AD-A031 238 TEXAS UNIV AT AUSTIN DEPT OF AEROSPACE ENGINEERING AN--ETC F/G 20/11
A STUDY OF CONVERGENCE AND STABILITY OF FINITE ELEMENT APPROXIM--ETC(U)
AUG 76 J T ODEN, L C WELLFORD, C T REDDY DAAG29-76-G-0022

UNCLASSIFIED ARO-11860.9-M NL

1 OF 4
ADA031238

A STUDY OF CONVERGENCE AND STABILITY OF FINITE ELEMENT APPROXIMATION OF SHOCK AND ACCELERATION WAVES IN NONLINEAR MATERIALS

J. T. Oden, L. C. Wellford, Jr., and C. T. Reddy

August, 1976

U. S. ARMY RESEARCH OFFICE

P-11860-M/DAAG29-76-G-0022

The University of Texas at Austin

A STUDY OF CONVERGENCE AND STABILITY OF FINITE ELEMENT
APPROXIMATION OF SHOCK AND ACCELERATION WAVES
IN NONLINEAR MATERIALS

J. T. Oden, L. C. Wellford, Jr., and C. T. Reddy

August, 1976

U. S. ARMY RESEARCH OFFICE

P-11860-M/DAAG29-76-G-0022

The University of Texas at Austin

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.

TABLE OF CONTENTS

CHAPTER IX

FORWARD

This document represents the final report for work done on Project DAAG29-76-G-0022, "A Study of Convergence and Stability of Finite Element Approximations of Shock and Acceleration Waves in Nonlinear Materials."

The duration of the research project was two years, and all work was conducted at the Texas Institute for Computational Mechanics in the Department of Aerospace Engineering and Engineering Mechanics at the University of Texas at Austin. Principal Investigator of the project was Professor J. Tinsley Oden and, during the course of the work he was assisted by a number of graduate students. These include L. C. Wellford, Jr., C. T. Reddy, N. Kikuchi, John O'Leary, Cid Gesteira, and K. Ohtake. The work has resulted in a number of papers and presentations. A complete list is given at the end of this forward.

The present volume summarizes some of the major ideas in connection with discontinuous finite element methods for the calculation of shocks. This report is compiled principally from some of the papers listed below, but mostly is extracted from an abridged version of L. C. Wellford's dissertation. Some of the error estimates contained in this report are unique to the study of Galerkin methods for nonlinear hyperbolic equations. Error estimates of this type were originally obtained by Oden and Wellford in the reports listed below, but they were based on standard error estimates for quasi-linear elliptic equations. The question of optimal error estimates for these problems con-

tinues to be open (for discussion of the basence of optimal error estimates in nonlinear elliptic equations, see the recent paper of Babuska, "Singularity Problems in the Finite Element Method," U.S./ German Symposium on Formulations and Computational Algorithms in Finite Element Analysis, M.I.T. Press, 1976). In recent months, Oden and eddy have uncovered techniques for improving these estimates, and these have led to a considerable improvement in the bounds obtained by Oden and Wellford. All of these new results have been incorporated into this final report.

J. T. Oden
Principal Investigator

Appendix A

Publications

1. Wellford, L. C., Jr. and Oden, J. T., "A Theory of Discontinuous Finite-Element Galerkin Approximations of Shock Waves in Nonlinear Elastic Solids, I. Variational Theory,", Computer Meth. Appl'd. Mech. and Engr'g., 7, No. 1, 1976.

2. Wellford, L. C., Jr. and Oden, J. T., "A Theory of Discontinuous Finite-Element Galerkin Approximations of Shock Waves in Nonlinear Elastic Solids, II. Accuracy and Convergence", Computer Meth. Appl'd. Mech. and Engr'g., 7, No. 2, 1976.

3. Wellford, L. C., Jr. and Oden, J. T., "Discontinuous Finite Element Approximations for the Analysis of Shock Waves in Nonlinearly Elastic Materials, J. Comp. Physics, 19, 1975, 179-210.

4. Oden, J. T. and Wellford, L. C., Jr., "Continuous and Discontinuous Finite Element Approximations of Shock Waves in Nonlinear Elastic Solids, Comp. Meth., Vol. 461, Springer-Verlag, Heidelberg, 1975, pp. 149-168.

5. Oden, J. T. and Wellford, L. C., Jr., "Finite Element Analysis of Shocks and Acceleration Waves in Nonlinearly Elastic Solids," Shock and Vib. Digest, Vol. 7, No. 2, Feb. 1975, pp. 2-11.

6. Oden, J. T. and Wellford, L. C., Jr., "Discontinuous Finite Element Approximations for the Analysis of Acceleration Waves in Elastic Solids," The Mathematics of Finite Elements with Applications, Brunel University, Uxbridge, England, April 1975 (to be published by Academic Press, London).

7. Wellford, L. C., Jr., "Variational Methods for Wave and Shock Propagation in Nonlinear Hyperelastic Materials," Proceedings, Inter. Conf. on Comp. Methods in Nonlinear Mech., Austin, Texas, Sept. 1974, pp. 879-889.

8. Wellford, L. C., Jr., "Finite Element Galerkin Methods for the Analysis of Wave and Shock Propagation in Hyperelastic Solids," Ph.D. Dissertation, The University of Alabama in Huntsville, 1974.

9. Oden, J. T. and Wellford, L. C., Jr., "Some New Finite Element Methods for the Analysis of Shock and Acceleration Waves in Nonlinear Materials, Finite Element Analysis of Transient Structural Behavior, ASME Winter Annual Meeting, Houston, Nov. 1975.

10. Oden, J. T., "Galerkin Approximations of a Class of Nonlinear Boundary-Value Problems and Evolution Problems in Elasticity," Second IRIA Symposium on the Computer Methods in Applied Science and Engineering, Paris, December 1975.

CHAPTER I

INTRODUCTION

I.1 Opening Comments. This report is concerned with the approximation of problems of wave and shock propagation in nonlinear solids through the finite-element implementation of the Galerkin method. For simplicity, we consider here only hyperelastic materials and one-dimensional domains. However, the methods to be developed are applicable to other materials and to more general domains.

Three computational schemes for wave and shock propagation are introduced. We choose to identify them as follows:

(i) Shock fitting method using discontinuous trial functions.

(ii) Parabolic regularization method.

(iii) A central difference method.

We note that the first two methods are valid for shock waves while the third method is useful for nonlinear dynamic response.

In this report we attempt to attain a proper balance between theory and calculation. On the theoretical side for each of the methods, we discuss the formulation, the convergence, the accuracy, and the stability of the approximation. On the computational side, we discuss the implementation of the methods and present numerical results.

I.1.1 The Shock Fitting Scheme. In the early 1950's, Von Neumann [1] conjectured that the problem of shock propagation through a given domain might be attacked by considering it as a free boundary type boundary-value problem for the partial differential equations involved--that is, a nonlinear boundary value problem in which the boundary varied with time and its location at any time was an unknown. Then the shock propagation problem could be characterized by a system of partial differential equations defined on certain shockless domains, together with a collection of boundary conditions (jump conditions) valid at the shock surfaces which form the interfaces of these domains.

For several reasons, not the least of which was a lack of a complete theory of free boundary value problems, the mathematical theory of shock waves has developed along quite different lines. The thrust of basic work in this area has been toward a global theory that would accomodate shocks, and this has necessarily led to viewing the shock solution in a distributional setting. That is, one seeks from the onset generalized solutions of certain nonlinear hyperbolic equations. For a recent account of the state of the subject, see Lax [2].

Our interest here in these interpretations is from the viewpoint of approximation; i.e., what specific types of approximation schemes do these contrasting philosophies suggest? The answer is fairly clear. The distributional theory forms the backbone of virtually all variational methods of approximation. In particular, it is well known that the Galerkin concepts, when used in conjunction with finite element interpolants, represent one of the most powerful methods of approximation of elliptic and parabolic equations. Among their attractive features are their accuracy, conditioning, and the ease with which they can be

applied to very irregular domains on which mixed boundary conditions are prescribed. Unfortunately, Galerkin schemes are not considered seriously for shock wave calculations. Increased continuity requirements are implied in the Galerkin formulation and, consequently, they endow the shock wave approximation with an unrealistic degree of smoothness. In other words, they tend to smooth solutions which are actually discontinuous, and, in transient problems, this often leads to an unacceptable amount of artificial dissipation. On the other hand, the free boundary conjectures of Von Neumann have been the basis of the shock-fitting schemes popular in gas dynamics calculations. There questions of irregular domains and complicated boundary conditions are rarely important, and the principal aim is to depict accurately the strength of the shock. It would seem that a scheme that adopted the advantages of both of these classes of methods would be the most appropriate for general shock problems.

The shock fitting scheme to be introduced here represents the implementation of these ideas using what we call a "generalized Galerkin" method. We extend the Galerkin method to handle shape functions with built in discontinuities. These discontinuities in the global basis are allowed to move through the Lagrangian finite element grid. In this way we generate a true free boundary formulation. The location of the free boundary (or shock) in this case becomes one of the dependent variables of the problem.

To construct the global finite-element model we use two types of finite element interpolating functions. One consists of the conventional piecewise polynomial approximations for representations over

shockless domains; the second involves the set of local trial functions with prescribed discontinuities that are introduced in a collection of elements forming a boundary layer around the shock surfaces. Then the "generalized Galerkin" method turns out to be a "local projectional" method in which the projections are taken on the shockless domains between moving boundaries (shocks).

In the course of the theoretical development, we show that, given the normal finite element interpolation property, the standard Galerkin method cannot converge to shock wave solutions. This is because the shock wave solution is very irregular globally and the standard Galerkin method is a global projection method. Then the obvious improvement offered by the shock fitting method is that it is a local projection method and the projections are taken on domains on which the solution is very regular. By using these ideas, we prove conclusively that the shock fitting scheme is able to retain the desirable properties of conventional Galerkin techniques (that is, accuracy, geometrical independence, stability) and at the same time model effectively the strength of the shock, its propagation and decay.

<u>I.1.2 The Parabolic Regularization Method</u>. In I.1.1, we commented that the nonconvergence of the standard Galerkin method to shock wave solution is caused by the irregular character of the solutions. The question is, can we somehow regularize the solutions and cause the convergence of the standard Galerkin procedure? The answer is 'yes'. We know from the Sobolev imbedding theorem [3] that there is an infinitely differentiable function which is very close to the shock wave solution (which has first temporal and spatial derivatives which are square

integrable). We can produce such a function by solving a regularized version of the original hyperbolic Galerkin model. We append to the original functional certain terms which make the problem parabolic rather than hyperbolic. We multiply these terms by small regularizing parameters depending on the discretization parameters. Then, as we shrink the discretization parameter to zero, the regularized solution approaches the original solution.

This procedure is quite similar to the shock smearing-artificial viscosity methods developed for finite difference calculations except that the viscosity depends somehow on the discretization parameters. It is well known that the artificial viscosity methods tend to dissipate the energy in the shock in order to obtain their regularizing objective. This damping is a major drawback in using the method.

In the parabolic regularization method developed here, we use the theoretical power inherent in the variational formulation to optimize the regularizing parameter through its dependence on the discretization parameters. We determine how to regularize the problem to insure convergence, accuracy, and stability and still minimize the dissipation.

<u>I.1.3 The Central Difference Method</u>. In order to obtain accurate numerical solutions for nonlinear dynamic response and wave propagation, a computationally efficient algorithm is desirable. A nonlinear central difference algorithm is proposed here. The accuracy and stability of the method are analyzed using L_2 methods. It is shown that the normal Courant-Friedricks-Lewy stability criteria [4] is required to obtain sufficient conditions for convergence. An additional stability

condition which limits the amplitude of the response is required for convergence in the nonlinear case. Essentially, we can develop sufficient conditions for convergence only for limited ranges of the size of the solution.

I.2 Historical Perspective. The initial application of approximation methods to shock waves was performed by Von Neumann [1] and Von Neumann and Richtmyer [5]. These applications were to problems in fluid dynamics and used what have become known as artificial viscosity methods. In these methods, terms are appended to the equations of motion to make them dissipative. In [1] and [5] these terms were quadratic in the velocity gradient. Landshoff [6] suggested a method which is linear instead of quadratic in the velocity gradients. Rusanov [7], on the other hand, has used an artificial viscosity method employing a diffusion type term involving the second spatial derivatives of the velocity.

The first of a series of implicit artificial damping methods was introduced by Lax [8]. Lax was able to implicitly introduce damping in difference equations by replacing certain velocity components by their averages in space. Lax and Wendroff [9,10] developed another family of implicit artificial damping methods by using a second order Taylor series expansion in time, Richtmyer and Morton [11] proposed an alternate and computationally more efficient version of this scheme using quantities evaluated at a half step forward in time. MacCormack [12] has introduced Lax-Wendroff type schemes for several space dimensions. Rusanov [13] has employed more accurate Lax-Wendroff schemes. Similar procedures have been introduced by Burstein and Mirin [14].

Another type of implicit artificial damping scheme has been introduced by Godunov [15]. This method used the exact solution to define certain fluxes in the conservative form of the equation. In some ways it is similar to the scheme of Glimm [16]. The technique of Courant, Isaacson and Rees [17] also involves implicit damping due to the one-sided differences involved.

Concerning the application of these difference type artificial viscosity methods to shock propagation in solids, we remark that the major work has been carried out at Sandia Laboratories and has been reviewed by Walsh [18]. The result of this research has been a series of one-dimensional computer codes using both quadratic and linear artificial viscosity (see Lawrence [19]). Recent refinements have included a rezoning procedure (see Lawrence and Mason [20]).

Concerning the application of noncharacteristic finite difference methods of the shock fitting type, the primary contributor has been Moretti (see [21,22,23,24,25,26,27,28]). Moretti's method uses a noncharacteristic moving mesh with characteristic calculations at the moving shock. Yu and Seebass [29] have presented a shock fitting method for transonic flow. Richtmyer and Morton (e.g. [11], section 13.9) have discussed certain aspects of two-dimensional shock fitting. Skoglund and Cole [30] have developed a 2D shock fitting method with a fixed grid. The implementation of this method seems difficult. Taylor [31] has introduced a shock fitting method involving local integration through the shock. Bellman, Cherry, and Wing [32] used a method which required following characteristics to integrate through the shock. They found that the method was unstable for all but Burger's equation.

Gary [33] used a shock fitting method for hydrodynamic problems containing shocks. Xerikos [34,35] introduced shock fitting methods in cylindrical coordinates. Weinbaum [36] has applied shock fitting in wake flow. We remark that there seem to have been no application of this type noncharacteristics method to shock propagation in solids.

We do not intend to fully review the literature on the method of characteristics for shock fitting. The applications of this technique to solids has been reviewed by Karpp and Chou [37]. In particular we note that the earliest numerical work seems to have been by Thomas [38]. More recent applications have been by Herrmann, Hicks, and Young [39] and Hicks and Holdridge [40] in the area of elastic-plastic stress waves.

The application of finite-element/Galerkin methods to shock propagation is a relatively new undertaking. We remark that Oden and Fost [41] have developed accuracy and convergence results for a central difference approximation for nonlinear one-dimensional elastodynamics problems. Fost [42] has applied a Lax-Wendroff type finite-element scheme to shock propagation in one and two-dimensional hyperelastic bodies. Fost, Oden, and Wellford [43] have analyzed finite-element methods for shock propagation in one-dimensional hyperelastic bodies and have presented numerical results. Oden, Key, and Fost [44] have discussed certain aspects of this method in the dynamics of membranes. Oden [45] has further developed the theoretical aspects of the approximation. In addition, Argyris, Dunne, Angelopoulos, and Bichat [46] have presented similar results using finite elements in time and artificial damping.

In the theoretical work concerning accuracy and convergence of finite-element/Galerkin methods we rely on the L_2 method for deriving error estimates. The basic paper concerning this method and its application to parabolic equations was contributed by Douglas and Dupont [47]. More recently Wheeler [48] refined this method and analyzed certain nonlinear problems. Dupont [49] has applied the L_2 method to the linear second order hyperbolic equations. In addition, Douglas and Dupont [50] have applied alternating direction methods for Galerkin approximations to hyperbolic equations. Dendy [51] has discussed some aspects of this method in application to nonlinear hyperbolic equations.

I.3 Objectives of the Report. The objectives of this report can be summarized as follows:

i) The development of generalized Galerkin methods using discontinuous trial functions to approximate problems with irregular solutions, together with application of the method to wave and shock propagation in solids.

ii) Implementation of the generalized Galerkin method using discontinuous finite elements. This effectively creates a finite element shock fitting technique for wave and shock propagation in solids.

iii) Calculation of the growth and decay of shock waves in solids using the generalized Galerkin shock fitting scheme of ii).

iv) Development of a theory of finite-element methods for a class of nonlinear static elasticity problems. Verification of the theory through numerical experiment.

v) Development of accuracy, convergence, and stability results for the generalized Galerkin shock fitting technique.

vi) Development of a parabolic regularization method for shock propagation in nonlinear solids. Study of the questions of accuracy, convergence, and stability of the method.

vii) Implementation of the parabolic regularization technique using finite element trial spaces.

viii) Development of a theory of central difference schemes for wave propagation in nonlinear solids.

In Chapter II, a class of static nonlinear elasticity problems is constructed. The stress is characterized mathematically. Continuity and strong monotone properties are determined.

In Chapter III, the accuracy and convergence properties of finite element approximations to these static nonlinear problems are analyzed. The results of numerical experiments to confirm these theories are discussed.

In Chapter IV, a theory of generalized Galerkin methods for problems with irregular solutions is presented. Applications are given to the problem of wave and shock propagation in nonlinear hyperelastic solids. Accuracy, convergence, and stability are analyzed.

In Chapter V, the generalized Galerkin method for shock propagation is implemented using discontinuous finite elements. The calculation of reflection of waves is discussed, and numerical re= sults are presented.

In Chapter VI, a theory of parabolic regularization methods for wave and shock propagation is constructed. The accuracy and convergence of the method are determined and a stability criteria is developed.

In Chapter VII, the parabolic regularization method is im-

plemented using finite element interpolants. Numerical results are presented.

In Chapter VIII, a theory of central difference approximations for wave propagation in nonlinear solids is presented.

In Chapter IX, conclusions are drawn and proposals for future extensions to this work are made.

C H A P T E R   I I

STATIC AND DYNAMIC BEHAVIOR OF NONLINEAR ELASTIC SOLIDS

II.1 Problem Classification. We wish to examine certain weak nonlinear elliptic boundary-value problems suggested by equations of the form [52]

$$\left.\begin{aligned} -D(\sigma(Du(X))) &= f(X),\ X \in I \\ u(0) = 0, u(a) &= 0 \end{aligned}\right\} \qquad (2.1.1)$$

We choose to interpret the quantities in (2.1.1) in the following way: $u(X)$ is the displacement component in the direction x of a particle $X(u = x - X)$ in a continuous body B subjected to body forces per unit initial volume of $f(X)$. Then I is a bounded open set of particles and $\{0,a\} \subset I$ is the closure of I. In (2.1.1), $D \equiv d/dX$ and the function $\sigma(Du)$ is the stress. Then (2.1.1) represents the local balance of linear momentum at X in the case of static or quasistatic motions of B; i.e. (2.1.1) is an equation of mechanical equilibrium. Moreover, the particular form of the function σ defines the mechanical constitutive properties of B.

We weaken conditions on u and the data f by considering, instead of (2.1.1), the following nonlinear variational problem: find $u \in \mathcal{U}$ such that

$$(\sigma(Du),\ Dv) = \langle f,v\rangle\ ,\quad \forall\ v \in \mathcal{U} \qquad (2.1.2)$$

Here $\mathcal{U}$ is a Banach space and σ is viewed as the extension of a nonlinear operator from $D(\mathcal{U})$ into some larger space F, containing $D(\mathcal{U})$ and with a weaker topology, in which the data is contained. The brackets $(\cdot,\cdot)$ and $\langle\cdot\ \cdot\rangle$ indicate duality pairings between $D(\mathcal{U})$ and its dual $(D(\mathcal{U}))'$ and $\mathcal{U}$ and $\mathcal{U}'$. We are more specific about the structure of these spaces below.

The character of solutions to (2.1.2), and, indeed, the existence of solutions to (2.1.2), depends upon the data and properties of function $\sigma(Du)$. We shall be concerned with materials in which $\sigma(Du)$ corresponds to an isotropic, hyperelastic material, in which case σ is derivable from a potential function $W(Du)$ which represents the strain energy per unit of initial volume V_0 of B, where V_0 is the volume of B when it occupies its reference configuration. Then we write

$$\sigma(Du) = \frac{\partial W(Du)}{\partial Du} \quad \text{or} \quad \sigma(\lambda) = \frac{\partial W(\lambda)}{\partial \lambda} \qquad (2.1.3)$$

where λ is the <u>extension ratio</u> or <u>stretch</u>,

$$\lambda(X) = 1 + Du(X) \qquad (2.1.4)$$

Although the functional forms $\sigma(Du)$ and $\sigma(\lambda)$ (or $W(Du)$ and $W(\lambda)$) are, of course, different, we shall use the same symbol σ (or W) for each, except when confusion is likely,

We remark that there are at least two nontrivial classes of nonlinear elastostatics problems corresponding to (2.1.2) and (2.1.3):

I. <u>Type I Problems</u>. For compressible materials, (2.1.2) and (2.1.3) describe the equilibrium of a finite slab of hyperelastic material,

$$B = \{\underset{\sim}{X}: \ \underset{\sim}{X} = (X,Y,Z) \in \mathbb{R}^3, \ 0 \le x \le a\}$$

in plane strain, on which body forces have been prescribed which are uniform in Y and Z, but vary with X. We include in this class the half space,

$$B = \{\underset{\sim}{X}: \ \underset{\sim}{X} = (X,Y,Z) \in \mathbb{R}^3, \ X \ge 0\}$$

wherein uniform displacements may be prescribed over the planes X = 0 and X = ∞.

II. Type II Problems: For incompressible materials, (2.1.2) may describe the plane, axisymmetric, longitudinal stretching of a thin cylindrical rod of hyperelastic material in which the principal stress components normal to the axis of the rod are taken to be zero. Then, instead of (2.1.3) we have

$$\sigma(\lambda,\mu(\lambda)) = \frac{\partial \hat{W}}{\partial \lambda} \ ; \ \hat{W} = W(\lambda,\mu) + h(\lambda^2\mu^4 - 1) \qquad (2.1.5)$$

where λ and $\mu = \lambda^{-1/2}$ are principal stretches and the hydrostatic pressure h is determined from the transverse stress condition $\partial\hat{W}/\partial\mu = 0$; i.e.

$$h = -\frac{1}{4\mu^3\lambda^2} \frac{\partial W(\lambda,\mu)}{\partial \mu} \ ; \ \mu = \lambda^{-1/2} \qquad (2.1.6)$$

We elaborate further on these classes of problems below.

II.2 Constitutive Characterization of Nonlinear Elastic Materials.

Throughout this section, the symbols σ and W will represent the stress and the strain energy, respectively, and the same symbols may be used to represent different forms of functions which depict the dependence of σ and W on various kinematical quantities.

For motions of isotropic hyperelastic materials of the type described previously, the strain energy can be expressed as a function of the principal invariants of the deformation tensor,

$$\left.\begin{aligned} I_1 &= \lambda^2 + \mu^2 + \nu^2 \\ I_2 &= \lambda^2\mu^2 + \lambda^2\nu^2 + \mu^2\nu^2 \\ I_3 &= \lambda^2\mu^2\nu^2 \end{aligned}\right\} \tag{2.2.1}$$

where λ, μ, and ν are the principal stretches. For the classes of problems described previously,

$$\lambda = \lambda(X) = 1 + Du(X); \; \mu = \nu = \mu(X) \tag{2.2.2}$$

Let the strain energy be given as a function $W(I_1,I_2,I_3)$ of the invariants I_i of (2.2.1). Then the stress σ in the direction X, and the normal stress s in a direction transverse to X and corresponding to the stretch μ are given by

$$\left.\begin{aligned} \sigma &= 2\lambda \frac{\partial W}{\partial I_1} + 2\lambda(\mu^2 + \nu^2) \frac{\partial W}{\partial I_2} + 2\lambda\mu^2\nu^2 \frac{\partial W}{\partial I_3} \\ s &= 2\mu \frac{\partial W}{\partial I_1} + 2\mu(\lambda^2 + \nu^2) \frac{\partial W}{\partial I_2} + 2\mu\lambda^2\nu^2 \frac{\partial W}{\partial I_3} \end{aligned}\right\} \tag{2.2.3}$$

We now restrict ourselves to the two classes of problems described previously.

I. Type I. In this case, the material is compressible, $\mu = \nu = 1$, $s \neq 0$, $W = W(I_1,I_2,I_3)$, and

$$\sigma(\lambda) = 2\lambda \left(\frac{\partial W}{\partial I_1} + 2 \frac{\partial W}{\partial I_2} + \frac{\partial W}{\partial I_3}\right) \tag{2.2.4}$$

II. Type II. In this case, the material is incompressible, $\mu = \nu = \lambda^{-1/2}$, and $2\partial W/\partial I_3$ is associated with the hydrostatic pressure h which is eliminated using the condition s = 0. Then $W = W(I_1, I_2)$ and

$$\sigma(\lambda) = 2(\lambda - \frac{1}{\lambda^2}) \frac{\partial W}{\partial I_1} + 2(1 - \frac{1}{\lambda^3}) \frac{\partial W}{\partial I_2} \qquad (2.2.5)$$

We shall further restrict our analysis to isotropic hyperelastic materials in which the strain energy is a polynomial in the principal invariants; i.e. we henceforth assume that W is of the form

$$W = \sum_{m=0}^{N} \sum_{p=0}^{N} \sum_{q=0}^{N} C_{mpq}(I_1 - 3)^m (I_2 - 3)^p (I_3 - 1)^q \qquad (2.2.6)$$

or, for incompressible materials,

$$W = \sum_{p=0}^{N} \sum_{q=0}^{N} C_{pq}(I_1 - 3)^p (I_2 - 3)^q \qquad (2.2.7)$$

where m, p, q, and N are nonnegative integers and C_{mpq} and C_{pq} are arrays of material constants. Introducing (2.2.6) and (2.2.7) into (2.2.4) and (2.2.5) respectively, we arrive at the following constitutive equations for Type I and Type II materials.

I.

$$\sigma(\lambda) = \sum_{0 \leq m,p,q, \leq N} \sum_{r=0}^{m+p+q-1} A_{mpqr} \lambda^{2(m+p+q-r)-1} \qquad (2.2.8)$$

$$\left.\begin{aligned} A_{mpqr} &= (-1)^r (2)^{p+2} C_{mpq}(m+p+q) \binom{m+p+q-1}{r} \\ A_{000r} &= 0 \quad \forall \quad 0 \leq r \leq 3N-1 \end{aligned}\right\} \qquad (2.2.9)$$

II.

$$\sigma(\lambda) = \sum_{\substack{p,q,r \\ s,t,k,e}} B_{pqrstke} \lambda^{2p+q-2r-3(s+k)-t-e}$$

$$+ \sum_{\substack{p,q,a \\ b,i,j,e}} D_{pqabije} \lambda^{2p+q-2a-3(b+j)-i-e} \tag{2.2.10}$$

$$B_{pqrstke} = p(-3)^{r+t} 2^{q+s-t-k-1} a_e \frac{(p-1)!}{r!s!(p-1-r-s)!}$$

$$\cdot \frac{q!}{t!k!(q-t-k)!} \cdot C_{pq}$$

$$0 \leq p,\ q \leq N;\ 0 \leq r \leq p-1;\ 0 \leq s \leq p-1-r$$

$$0 \leq t \leq q;\ 0 \leq k \leq q-t;\ 1 \leq e \leq 4 \tag{2.2.11}$$

$$a_1 = -a_4 = 1,\ a_2 = a_3 = 0$$

$$D_{pqabije} = q(-3)^{a+i}\ 2^{b+q-i-j}\ a_e \frac{p!}{a!b!(p-a-b)!}$$

$$\frac{(q-1)!}{i!j!(q-1-i-j)!} \cdot C_{pq}$$

$$0 \leq a \leq p;\ 0 \leq b \leq p-a$$

$$0 \leq i \leq q-1;\ 0 \leq j \leq q-1-i$$

Thus, we see that for problems of both Types I and II, we can write

$$\sigma(\lambda) = \sum_{k=M_1}^{M_2} C_k \lambda^k \tag{2.2.12}$$

where for Type I, $M_1 > 0$ and for Type II, $M_1 < 0$, M_1 and M_2 being integers.

Example: We mention, as examples, the way in which several of the forms of $W(I_1,I_2)$ that have been proposed for incompressible rubbery materials lead to stress-strain laws of the form (2.2.7). First, the neo-Hookean material (e.g. [53]), for which

$$W = C_{10}(I_1 - 3) \tag{2.2.13}$$

Then the well-known Mooney form [54], proposed for certain natural rubbers,

$$W = C_{10}(I_1 - 3) + C_{01}(I_2 - 3) \tag{2.2.14}$$

and, as an example of a higher-order polynomial form, we mention a class of materials described by Biderman [55]

$$W = C_{10}(I_1 - 3) + C_{20}(I_1 - 3)^2 + C_{30}(I_1 - 3)^3 + C_{01}(I_2 - 3) \tag{2.2.15}$$

and a form suggested by Isihari, Hatshitsumi, and Tatibana [56] on the basis of a kinetic theory for rubbery materials:

$$W = C_{10}(I_1 - 3) + C_{20}(I_2 - 3)^2 + C_{01}(I_2 - 3) \tag{2.2.16}$$

Introducing each of these into (2.2.10) we find that the constitutive laws assume the form (2.2.12) with the following values of the constants M_1, M_2, and C_k:

Neo-Hookean

$$\left.\begin{array}{l} M_1 = -2,\ M_2 = 1 \\ C_{-2} = 2C_{10},\quad C_1 = 2C_{10},\quad \text{all other } C_k = 0 \end{array}\right\} \tag{2.2.17}$$

Mooney

$$\left.\begin{aligned}&M_1 = -3,\ M_2 = 1\\ &C_{-3} = -2C_{01},\quad C_{-2} = -2C_{10},\quad C_0 = 2C_{01}\\ &C_1 = 2C_{10},\quad \text{all other } C_k = 0\end{aligned}\right\} \tag{2.2.18}$$

Biderman

$$\begin{aligned}&M_1 = -3,\ M_2 = 5\\ &\left.\begin{aligned}&C_{-3} = 48C_{30} - 8C_{20} - 2C_{01}\\ &C_{-2} = 12C_{20} - 2C_{10} - 54C_{30}\\ &C_{-1} = -24C_{30};\quad C_0 = 2C_{01} + 4C_{20}\\ &C_1 = 2C_{10} + 54C_{30} - 12C_{20}\\ &C_2 = 18C_{30};\ C_3 = 4C_{20} - 3C_{30};\ C_4 = 0\end{aligned}\right\}\\ &C_5 = 6C_{30}\end{aligned} \tag{2.2.19}$$

IHT

$$\left.\begin{aligned}&M_1 = -4,\ M_2 = 2\\ &C_{-4} = -4C_{20};\quad C_{-3} = -2C_{01};\quad C_{-2} = -2C_{10}\\ &C_{-1} = 2C_{10} - 4C_{20};\quad C_0 = 2C_{01}\\ &C_1 = 2C_{10};\quad C_2 = 8C_{10}\end{aligned}\right\} \tag{2.2.20}$$

The classes of materials and kinematical restrictions described above suggest a number of specific physically reasonable material properties. We observe that, in general, the following properties may be assumed to hold:

P.1 (Convexity). The strain energy function $W(\lambda)$ is a convex function of λ.

P.2 (Polynomial Boundedness from Below). There exists an integer $\alpha > 0$ and a positive constant $K_1 > 0$ such that

$$|\sigma(\lambda)| \geq K_1\lambda^{\alpha} \tag{2.2.21}$$

for all $\lambda \in (0,\infty)$.

P.3 (Almost Polynomial Growth). There exists a real-valued convex function b: $\mathbb{R} \to \mathbb{R}$ such that

$$\lim_{\lambda\to\infty} b(\lambda) = 0 \tag{2.2.22}$$

$$\lim_{\lambda\to 0} b(\lambda) = \infty \tag{2.2.23}$$

and there exists a positive constant $K_2 > 0$ such that

$$|\sigma(\lambda)| \leq b(\lambda) + K_2\lambda^{\alpha} \tag{2.2.24}$$

where α is an integer > 0.

Property P.2 simply insures that $\sigma(\lambda)$ is bounded below by a monotonic function of λ; then $\sigma(\lambda) \to \infty$ as $\lambda \to \infty$. Property P.3 asserts that within a real-valued function $b(\lambda)$, $\sigma(\lambda)$ grows slower than a polynomial in λ as λ tends to infinity; also, it implies that no finite volume of material can be shrunk to zero as $\lambda \to 0$. It can be verified that the examples (2.2.13) - (2.2.16) satisfy these hypotheses when the coefficients C_k in (2.2.12) are selected as indicated.

II.3. Notation and Definitions. We employ the following notations: $W_p^m(I)$ shall denote the Sobolev space of order m,p of functions defined on $I \subset \mathbb{R}$ whose generalized derivatives of order $\leq$ m are in the Lebesgue space $L_p(I)$, $1 \leq p < \infty$. It can be shown that for domains such

as I, $W_p^m(I)$ coincides with the completion of the space $C^m(I)$ of m-times continuously differentiable functions on I in the m,p-Sobolev norm

$$\|u\|_{W_p^m(I)} = \left\{ \int_I \sum_{k\leq m} |D^k u|^p \, dX \right\}^{1/p} \tag{2.3.1}$$

In the analysis that follows, we use the seminorm

$$\|u\|_{W_p^1(I)} = \left(\int_I |Du|^p \, dx \right)^{1/p} \tag{2.3.2}$$

The spaces $W_p^m(I)$ are complete, and the spaces $W_2^m(I)$ are Hilbert spaces. We use the notation

$$W_2^m(I) = H^m(I) \tag{2.3.3}$$

The inner product on $H^m(I)$ is given by

$$(u,v)_m = \int_I \sum_{k\leq m} D^k u \, \overline{D^k v} \, dX; \quad u,v \in H^m(I) \tag{2.3.4}$$

We also denote by $\mathring{W}_p^m(I)$ the completion of $C_o^m(I)$ in the norm (2.3.1) where $C_o^m(I)$ is the space of C^m-functions with compact support in I. Then $\mathring{W}_2^m(I) \equiv H_o^m(I)$. It is important to note the inclusion properties

$$W_p^m(I) \subseteq W_q^r(I); \quad m \geq r, \quad p \geq q \tag{2.3.5}$$

and, in particular,

$$W_p^m(I) \subseteq H^m(I), \quad p \geq 2 \tag{2.3.6}$$

Upon examining (2.1.2) and (2.2.8) or (2.2.10), it appears that the displacement $u(X)$ at least has first derivatives in $L_2(I)$

whenever the data $f(X,t) \in L_2(I)$ for each t. Since v(X) in (2.1.2) can be selected to have locally integrable first derivatives, we argue that for $f \in L_2(I)$, $u \in H^1(I)$. This is a Hilbert space. Of course, u also belongs to a smaller Banach space $\mathcal{U}$ of functions with finite energy, and v is selected from $\mathcal{U}$. Since we may admit distributional solutions, f need not be in $L_2(I)$, but in any case, we take $u \in \mathcal{U} \subset H^1(I)$ where, for problems of Type I and II, $\mathcal{U} = W^1_{p+1}(I)$ for some $p \geq 1$.

These considerations prompt us to adopt the following setting for our study. Let I again be an open interval of $\mathbf{R}$ and V and H be two Hilbert spaces of functions defined on I with $V \subset H$. Let $\mathcal{U}$ denote a Banach space with $\mathcal{U} \subset H$ and denote V' and $\mathcal{U}'$ the topological duals of V and $\mathcal{U}$. The stress σ and the strain energy W are regarded as mappings of the type σ: $\mathcal{D}(V) \cap \mathcal{D}(\mathcal{U}) \to \mathcal{D}(V)' + \mathcal{D}(\mathcal{U})'$; $W: \mathcal{D}(V) \cap \mathcal{D}(\mathcal{U}) \to \mathbf{R}$.

We recall that an operator $\sigma: \mathcal{D}(V) \cap \mathcal{D}(\mathcal{U}) \cap \mathcal{D}(V)' + \mathcal{D}(\mathcal{U})'$ is <u>monotone</u> on $\mathcal{D}(V) \cap \mathcal{D}(\mathcal{U})$ if and only if

$$(\sigma(u) - \sigma(v), u - v) \geq 0 \quad \forall\, u,v \in \mathcal{D}(V) \cap \mathcal{D}(\mathcal{U}) \tag{2.3.7}$$

where $(\cdot,\cdot)$ denotes the duality pairing on $\mathcal{D}(\mathcal{U})$ and $\mathcal{D}(\mathcal{U})'$. Moreover, σ is <u>strongly monotone</u> on $\mathcal{D}(V) \cap \mathcal{D}(\mathcal{U})$ if there exists a constant $\gamma > 0$ and a natural number r such that

$$(\sigma(u) - \sigma(v), u - v) \geq \gamma \, \|u - v\|^r_{\mathcal{D}(\mathcal{U})} \quad \forall\, u,v \in \mathcal{D}(V) \cap \mathcal{D}(\mathcal{U}) \tag{2.3.8}$$

Now, the Gateaux derivative will serve as a powerful tool in determining the properties of the stress σ. Let $u \in \mathcal{D}(V) \cap \mathcal{D}(\mathcal{U})$ and f be a function on $\mathcal{D}(V) \cap \mathcal{D}(\mathcal{U})$. We say that f is <u>Gateaux differentiable</u> at point $u \in \mathcal{D}(V) \in \mathcal{D}(\mathcal{U})$ if for all $\eta \in \mathcal{D}(V) \cap \mathcal{D}(\mathcal{U})$, the

mapping $\alpha \to f(u + \alpha\eta)$ is differentiable at $\alpha = 0$. Then the Gateaux derivative $D_G f(u)\ (\eta)$ at u in direction η is given by

$$D_G f(u)\ (\eta) = \lim_{\alpha\to 0} \left[\frac{f(u+\alpha\eta) - f(u)}{\alpha}\right] \qquad (2.3.9)$$

II.4 Monotone Behavior of σ. Now the concept of monotoncity of σ and convexity of W are related through the following theorem (Kachurovskii [57]).

Theorem 2.1. Let W: $\mathcal{D}(V) \cap \mathcal{D}(U) \to \mathbb{R}$ be differentiable in the sense of Gateaux. Then the following conditions are equivalent:

i) W is convex

ii) The operator $\sigma = \frac{\partial W(u)}{\partial u}$ is monotone

Proof. Assume W is convex. Define an auxiliary function χ_α by

$$\chi_\alpha = W(\alpha u + (1-\alpha)v) \qquad \alpha \in [0,1]$$

$$u,v \in \mathcal{D}(V) \cap \mathcal{D}(U) \qquad (2.4.1)$$

Now take the Gateaux derivative of χ_α

$$D_G \chi_\alpha = \lim_{\alpha\to 0} \frac{W(2\alpha u + (1-2\alpha)v) - W(\alpha u + (1-\alpha)v)}{\alpha} \qquad (2.4.2)$$

Since W is convex

$$W(\alpha u + (1-\alpha)v) \le \alpha W(u) + (1-\alpha)W(v) \quad \alpha \in [0,1]$$

$$u,v \in \mathcal{D}(V) \cap \mathcal{D}(U) \qquad (2.4.3)$$

Thus

$$D_G\chi_\alpha \le \lim_{\alpha\to 0} \frac{2\alpha W(u) + (1-2\alpha)W(v) - W(\alpha u + (1-\alpha)v)}{\alpha} \tag{2.4.4}$$

and

$$D_G\chi_\alpha \ge \lim_{\alpha\to 0} \frac{W(2\alpha u + (1-2\alpha)v) - \alpha W(u) - (1-\alpha)W(v)}{\alpha} \tag{2.4.5}$$

But since $W(u)$ is convex, $D_G\chi_\alpha$ is increasing and $D_G\chi_o \le D_G\chi_1$. Using (2.4.4) evaluated at $\alpha = 1$ and (2.4.5) evaluated at $\alpha = 0$

$$(\sigma(u) - \sigma(v), u-v) = (D_G W(u) - D_G W(v), u-v) \ge D_G\chi_1 = D_G\chi_o \ge 0$$

Thus $\sigma(u)$ is a monotone operator.

Conversely, assume $\sigma(u)$ is a monotone operator. Then $W(u)$ is differentiable and increasing. Thus $\chi(\sigma)$ is convex and W is likewise. ■

Thus the obvious initial mathematical hypothesis following from P.1 - P.3 are

M.1 The stress σ is Gateaux differentiable.

M.2 The stress σ is a monotone operator.

As an example consider the Mooney material. From (2.2.18) we see that

$$\frac{\partial\sigma(\lambda)}{\partial\lambda} = 2C_{10} + \frac{4C_{10}}{\lambda^3} + 6\frac{C_{01}}{\lambda^4} \tag{2.4.6}$$

Clearly W is convex for positive values of C_{01} or C_{10}. (We note that the positive signs are the only permissible choice physically since the material is elastic and must return to a stress-free state at $\lambda = 1$). Now from (2.2.1) and (2.2.5)

$$W(u+\alpha\eta) = C_1\{ \frac{2}{u+\alpha\eta} + (u+\alpha\eta)^2 - 3\}$$

$$+ C_2 \{2(u+\alpha\eta) + \frac{1}{(u+\alpha\eta)^2} - 3\}$$

Thus

$$W(u+\alpha\eta) = C_1 \{\frac{-2\alpha\eta}{u(u+\alpha\eta)} + (2\alpha u\eta+\alpha^2\eta^2)\}$$

$$+ C_2 \{2\alpha\eta - \frac{2\alpha u\eta+\alpha^2\eta^2}{u^2(u^2+2\alpha u\eta+\alpha^2\eta^2)} \}$$

and using (2.3.8)

$$D_G W(u)(\eta) = C_1\{- \frac{\eta}{u^2} + 2u\eta\} + C_2 \{2\eta - \frac{2\eta}{u^3}\}$$

Thus the strain energy function W for the Mooney material is Gateaux differentiable. Theorem 2.1 then implies that $\sigma(\lambda)$ is a monotone operator. Similar results can be obtained from the other materials introduced in Section II.2.

II.5 Continuity Properties of σ. We wish to consider here materials which have a continuous first Piola-Kirchhoff stress tensor in an appropriate norm. We introduce this as our third mathematical hypothesis.

M.3 The stress σ is continuous.

We consider below several examples chosen from the materials presented in Section II.2.

We begin with problems of Type I for which $\sigma(\lambda) = \sigma(1 + Du)$ is given by (2.2.9) or (2.2.12) with $M_1 > 0$ and $M_1 = p$ = integer > 1.

Theorem 2.2. Let the stress be given as a polynomial in λ of the form (2.2.12) with $M_1 = 0$ and $M_2 = p$, where p is an integer > 1. Let $\bar{\lambda}, \lambda \in L_s(I)$, $s \geq p+1$. Then

$$\| \sigma(\bar{\lambda}) - \sigma(\lambda) \|_{L_q(I)} \leq g(\bar{\lambda},\lambda) \; \| \bar{\lambda} - \lambda \|_{L_{p+1}(I)} \tag{2.5.1}$$

where

$$q = 1 + \frac{1}{p} \tag{2.5.2}$$

and $g(\bar{\lambda},\lambda)$ is a nonnegative function on $\mathbf{R}^2 = (0,\infty) \times (0,\infty)$ given by

$$g(\bar{\lambda},\lambda) = C \sum_{k=0}^{p-1} \| \bar{\lambda} + \lambda \|^k_{L_k(\frac{p+1}{p-1})(I)} \tag{2.5.3}$$

in which

$$C = \max_{0 \leq j \leq p} |C_j|$$

and C_j are the material constants appearing in (2.2.12).

Proof. From (2.2.12) we observe that

$$\sigma(\bar{\lambda}) - \sigma(\lambda) = \sum_{k=0}^{p} C_k(\bar{\lambda}^k - \lambda^k) = \sum_{k=1}^{p} C_k(\bar{\lambda}^k - \lambda^k)$$

$$\leq \hat{C}|\bar{\lambda} - \lambda| \sum_{k=0}^{p-1} (\bar{\lambda} + \lambda)^k$$

where $\hat{C} = \max_j |C_j|$.

We have,

$$\langle \sigma(\bar{\lambda}) - \sigma(\lambda), w \rangle \leq \hat{C} \int_I w|\bar{\lambda} - \lambda| \sum_{k=0}^{p-1} (\bar{\lambda}^k + \lambda)^k \, dx \tag{2.5.4}$$

where $\langle\cdot,\cdot\rangle$ denotes a duality pairing on $L_q - L_p$.

Next consider a typical term of (2.5.4) and recall the multi-function Hölder's inequality

$$\int_I u_1 u_2 \cdots u_n \, dx \le ||u||_{L_{t_1}(I)} ||u||_{L_{t_2}(I)} \cdots ||u||_{L_{t_n}(I)}$$

$$\frac{1}{t_1} + \frac{1}{t_2} + \cdots + \frac{1}{t_n} = 1$$

with $t_1 = t_2 = p+1$, $t_3 = t_4 = \cdots = t_n$, $0 < k \le p-1$, we have

$$\int_I w|\bar{\lambda} - \lambda|(\bar{\lambda} + \lambda)^k \, dx \le ||w||_{L_{p+1}(I)} ||\bar{\lambda} - \lambda||_{L_{p+1}(I)} ||\bar{\lambda} + \lambda||^k_{k(\frac{p+1}{p-1})}$$

for $k = 0$, we recall the elementary inequality

$$||u||_{L_r(I)} \le L^{\frac{1}{r} - \frac{1}{s}} ||u||_{L_s(I)} \qquad r \le s \tag{2.5.5}$$

where $L = \mathrm{mes}(I)$. Then,

$$\int_I w|\bar{\lambda} - \lambda| dx \le ||w||_{L_{p+1}(I)} ||\bar{\lambda} - \lambda||_{L_{\frac{p+1}{p}}(I)}$$

$$\le C_p ||w||_{L_{p+1}(I)} ||\bar{\lambda} - \lambda||_{L_{p+1}(I)}$$

where $C_p = L^{\frac{p-1}{p+1}}$. Therefore

$$\langle\sigma(\bar{\lambda}) - \sigma(\lambda), w\rangle \le \hat{C} ||w||_{L_{p+1}(I)} ||\bar{\lambda} - \lambda||_{L_{p+1}(I)} [C_p + \sum_{k=1}^{p-1} ||\bar{\lambda} + \lambda||^k_{L_{k(\frac{p+1}{p-1})}(I)}]$$

Denoting $C = \hat{C}(\max(1, C_p))$,

$$\langle\sigma(\bar{\lambda}) - \sigma(\lambda), w\rangle \leq C||w||_{L_{p+1}(I)} ||\bar{\lambda} - \lambda||_{L_{p+1}(I)} \left(\sum_{k=0}^{p-1} ||\bar{\lambda} + \lambda||^k_{L_{k(\frac{p+1}{p-1})}(I)}\right) \tag{2.5.6}$$

Let q be given by (2.5.2). Then

$$||\sigma(\lambda) = \sigma(\lambda)||_{L_q(I)} = \sup_{w \neq 0} \left\{ \frac{|(\sigma(\bar{\lambda}) - \sigma(\lambda), w)|}{||w||_{L_{p+1}(I)}} \right\} \tag{2.5.7}$$

Thus, introducing (2.5.7) into (2.5.6) we get

$$||\sigma(\bar{\lambda}) - \sigma(\lambda)||_{L_q(I)} \leq \{C \sum_{k=0}^{p-1} ||\bar{\lambda} + \lambda||^k_{L_{k(\frac{p+1}{p-1})}(I)}\} ||\lambda - \lambda||_{L_{p+1}(I)}$$

To complete the proof we need only denote $g(\bar{\lambda},\lambda)$, the term in braces in the above equation.∎

Now, when $\lambda \in L_{p+1}(I)$, $u \in W^1_{p+1}(I)$, because $Du = \lambda - 1$. Thus, we also have the following :

<u>Corollary 2.1</u>. Let $\sigma(Du)$ be given by (2.2.13) with $\lambda = 1 + Du$, $M_1 = 0$, $M_2 = p > 1$, $\bar{u}, u \in W^1_s(I)$, $s \geq p+1$. Then

$$||\sigma(D\bar{u}) - \sigma(Du)||_{L_q(I)} \leq g(1 + D\bar{u}, 1 + Du)\ ||\bar{u} - u||_{W^1_{p+1}(I)} \tag{2.5.8}$$

where $q = (p+1)/p$ and $g(x,y)$ is defined by (2.5.3).∎

Passing on to Type II problems, we can derive results similar to (2.5.1) for certain specific types of incompressible materials. As examples, we shall prove some results for specific materials.

<u>Theorem 2.3</u>. Let $\bar{\lambda}$, $\lambda \in L_p(I)$, $p \geq 2$, and let $\sigma(\lambda)$ be given by the Mooney form; i.e. (2.2.13) with the choice of coefficients described in (2.2.18). Then

$$\| \sigma(\bar{\lambda}) - \sigma(\lambda) \|_{L_2(I)} \le g_1(\bar{\lambda},\lambda) \; \| \bar{\lambda} - \lambda \|_{L_2(I)} \tag{2.5.9}$$

where $g_1(\bar{\lambda},\lambda)$ is the nonnegative function

$$g_1(\bar{\lambda},\lambda) = 2C_{01} \; \| \frac{(\bar{\lambda}^2 + \bar{\lambda}\lambda + \lambda^2)}{\bar{\lambda}^3\lambda^3} \|_{L_\infty(I)} + 2C_{10} \; [\| \frac{\bar{\lambda} + \lambda}{\bar{\lambda}^2\lambda^2} \|_{L_\infty(I)} + 1] \tag{2.5.10}$$

Proof.

$$\sigma(\bar{\lambda}) - \sigma(\lambda) = (-\frac{2C_{01}}{\bar{\lambda}^3} - \frac{2C_{10}}{\bar{\lambda}^2} + 2C_{01} + 2C_{10}\lambda)$$

$$- (- \frac{2C_{01}}{\lambda^3} - \frac{2C_{10}}{\bar{\lambda}^2} + 2C_{01} + 2C_{1C}\lambda)$$

or

$$\sigma(\bar{\lambda}) - \sigma(\lambda) = 2C_{01}(\frac{\bar{\lambda}^3 - \lambda^3}{\bar{\lambda}^3\lambda^3}) + 2C_{10}(\frac{\bar{\lambda}^2 - \lambda^2}{\bar{\lambda}^2\lambda^2}) + 2C_{10}(\bar{\lambda} - \lambda)$$

$$= 2C_{01} \frac{(\bar{\lambda} - \lambda)(\bar{\lambda}^2 + \bar{\lambda}\lambda + \lambda^2)}{\bar{\lambda}^3\lambda^3} + 2C_{10} \frac{(\bar{\lambda} - \lambda)(\bar{\lambda} + \lambda)}{\bar{\lambda}^2\lambda^2}$$

$$+ 2C_{10}(\bar{\lambda} - \lambda)$$

Thus,

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), w) = 2C_{01}(\frac{(\bar{\lambda} - \lambda)(\bar{\lambda}^2 + \bar{\lambda}\lambda + \lambda^2)}{\bar{\lambda}^3\lambda^3}, w)$$

$$+ 2C_{10}(\frac{(\bar{\lambda} - \lambda)(\bar{\lambda} + \lambda)}{\bar{\lambda}^2\lambda^2}, w)$$

$$+ 2C_{10}(\bar{\lambda} - \lambda, w)$$

Using the Hölder inequality with $p=1$, $q=\infty$ and then the Schwarz inequality we obtain

$$(\langle\sigma(\lambda) - \sigma(\lambda),w\rangle) \leq 2C_{01}\|\bar{\lambda} - \lambda\|_{L_2(I)} \|w\|_{L_2(I)} \left\| \frac{\bar{\lambda}^2 + \bar{\lambda}\lambda + \lambda^2}{\bar{\lambda}^3\lambda^3} \right\|_{L_\infty(I)}$$

$$+ 2C_{10}\|\bar{\lambda} - \lambda\|_{L_2(I)} \|w\|_{L_2(I)} \left\| \frac{\bar{\lambda} + \lambda}{\bar{\lambda}^2\lambda^2} \right\|_{L_\infty(I)}$$

$$+ 2C_{10}\|\bar{\lambda} - \lambda\|_{L_2(I)} \|w\|_{L_2(I)}$$

which, in conjunction with the definition of the operator norm, leads to (2.5.9) and (2.5.10).■

As another example, consider the Biderman material:

Theorem 2.4. Let $\bar{\lambda}$, $\lambda \in L_p(I)$, $p \geq 6$, and let $\sigma(\lambda)$ be given by the Biderman form; i.e., (2.2.12) with choice of coefficients displayed in (2.2.19). Then

$$\|\sigma(\bar{\lambda}) - \sigma(\lambda)\|_{L_{\frac{6}{5}}(I)} \leq g_2(\bar{\lambda},\lambda)\ \|\bar{\lambda} - \lambda\|_{L_6(I)} \tag{2.5.11}$$

where $g_2(\bar{\lambda},\lambda)$ is the nonnegative function on $(0,\infty) \times (0,\infty)$ given by

$$g_2(\bar{\lambda},\lambda) = |C_{-3}|\ \left\| \frac{(\bar{\lambda} + \lambda)^3}{\bar{\lambda}^4\lambda^4} \right\|_{L_{\frac{3}{4}}(I)} + |C_{-2}|\ \left\| \frac{(\bar{\lambda} + \lambda)^2}{\bar{\lambda}^3\lambda^3} \right\|_{L_{\frac{3}{2}}(I)}$$

$$+ \ldots + |C_9|\ \|\bar{\lambda} + \lambda\|^4_{L_6(I)} \tag{2.5.12}$$

Proof. Since

$$\|\sigma(\bar\lambda) - \sigma(\lambda)\|_{L_{\frac{6}{5}}(I)} = \sup\left\{\frac{|(\sigma(\bar\lambda) - \sigma(\lambda),w)|}{\|w\|_{L_6(I)}} ;\ w \neq 0\right\}$$

we compute

$$\sigma(\bar\lambda) - \sigma(\lambda) = C_{-3}\,\frac{(\bar\lambda - \lambda)(\bar\lambda^3 + \bar\lambda^2\lambda + \bar\lambda\lambda^2 + \lambda^3)}{\bar\lambda^4\lambda^4}$$

$$+ C_{-2}\,\frac{(\bar\lambda - \lambda)(\bar\lambda^2 + \bar\lambda\lambda + \lambda^2)}{\bar\lambda^3\lambda^3} + \ldots$$

$$+ C_9(\bar\lambda - \lambda)(\bar\lambda^4 + \bar\lambda^3\lambda + \bar\lambda^2\lambda^2 + \bar\lambda\lambda^3 + \lambda^4)$$

$$\leq C_{-3}\,\frac{(\bar\lambda - \lambda)(\bar\lambda + \lambda)^3}{\bar\lambda^4\lambda^4} + C_{-2}\,\frac{(\bar\lambda - \lambda)(\bar\lambda + \lambda)^2}{\bar\lambda^3\lambda^3}$$

$$+ \ldots + C_9(\bar\lambda - \lambda)(\bar\lambda + \lambda)^4$$

This implies that

$$(\sigma(\bar\lambda) - \sigma(\lambda),w) \leq C_{-3}\left(\frac{(\bar\lambda - \lambda)(\bar\lambda + \lambda)^3}{\bar\lambda^4\lambda^4},w\right) + C_{-2}\left(\frac{(\bar\lambda - \lambda)(\bar\lambda + \lambda)^2}{\bar\lambda^3\lambda^3},w\right)$$

$$+ \ldots + C_9((\bar\lambda - \lambda)(\bar\lambda + \lambda)^4,w)$$

Using the Hölder inequality with p=3, q=3/2, we get

$$(\sigma(\bar\lambda) - \sigma(\lambda),w) \leq |C_{-3}|\ \|(\bar\lambda - \lambda)w\|_{L_{\frac{6}{2}}(I)}\ \left\|\frac{(\bar\lambda + \lambda)^3}{\bar\lambda^4\lambda^4}\right\|_{L_{\frac{3}{2}}(I)}$$

$$+ |C_{-2}|\ \|(\bar\lambda - \lambda)w\|_{L_{\frac{6}{2}}(I)}\ \left\|\frac{(\bar\lambda + \lambda)^2}{\bar\lambda^3\lambda^3}\right\|_{L_{\frac{3}{2}}(I)}$$

$$+ \ldots + |C_9|\ \|(\bar\lambda - \lambda)w\|_{L_{\frac{6}{2}}(I)}\ \|(\bar\lambda + \lambda)\|^4_{L_6(I)}$$

Then using the Schwarz inequality;

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), w) \leq \{ |C_{-3}| \; \| \frac{(\bar{\lambda} + \lambda)^3}{\lambda^4 \bar{\lambda}^4} \|_{L_{\frac{3}{2}}(I)}$$

$$+ \ . \ . \ . + |C_9| \; \| \bar{\lambda} + \lambda \|^4_{L_6(I)} \} \; \| \bar{\lambda} - \lambda \|_{L_6(I)} \; \| w \|_{L_6(I)}$$

The assertions of the theorem now follow from (2.5.12).∎

<u>Corollary 2.2</u>. Let the conditions of Theorem 2.3 hold. Then

$$\| \sigma(D\bar{u}) - \sigma(Du) \|_{L_2(I)} \leq g_1(1 + D\bar{u}, 1 + Du) \; \| \bar{u} - u \|_{W_2^1(I)} \tag{2.5.13}$$

where $g_1(\cdot,\cdot)$ is given by (2.5.10).∎

<u>Corollary 2.3</u>. Let the conditions of Theorem 2.4 hold. Then

$$\| \sigma(D\bar{u}) - \sigma(Du) \|_{L_{\frac{6}{5}}(I)} \leq g_2(1 + D\bar{u}, 1 + Du) \; \| \bar{u} - u \|_{W_6^1(I)} \tag{2.5.14}$$

where $g_2(\cdot,\cdot)$ is given by (2.5.12).∎

<u>II.6 Strongly Monotone Property of σ</u>. We wish to consider here materials which in addition to the properties introduced previously have the strong monotoncity property given by (2.3.7). We introduce this property as our fourth mathematical hypothesis

M.4 The stress σ is strongly monotone.

We now establish a number of theorems which show that the stress operators discussed in Section II.2 have strongly monotone

behavior. We use here a method of proof similar to one developed by Glowinski and Marrocco [58,59].

We introduce two important lemmas.

Lemma 2.1. Let x and y be positive real numbers and $k \in Z^+$. Then

$$(x^k - y^k)(x - y) \geq (\tfrac{1}{2})^k \; |x - y|^{k+1} \qquad (2.6.1)$$

Proof.

$$(x^k - y^k)(x - y) = x^{k+1} + y^{k+1} - xy(x^{k-1} + y^{k-1})$$

Note that $2xy = x^2 + y^2 - |x - y|^2$. We have,

$$(x^k - y^k)(x - y) = \frac{1}{2}[\,(x^{k-1} - y^{k-1})(x^2 - y^2) + (x^{k-1} + y^{k-1})|x - y|^2]$$

$$\geq \frac{1}{2}\,\frac{x^{k-1} + y^{k-1}}{|x - y|^{k-1}}\;|x - y|^{k+1} \geq \frac{1}{2}\left[\frac{(x^{k-1} + y^{k-1})^{\frac{1}{k-1}}}{(x + y)}\right]^{k-1} |x - y|^{k+1}$$

From the inequality $\dfrac{(x^p + y^p)^{1/p}}{x + y} \geq \dfrac{1}{2}$ the result follows immediately.■

Lemma 2.2. $\lambda, \bar{\lambda} > 0,\ k \in Z^+$

$$\int_I -(\frac{1}{\lambda^k} - \frac{1}{\lambda^k})\,(\bar{\lambda} - \lambda)dx \geq (\tfrac{1}{2})^k\; g_k(\lambda,\bar{\lambda})\,||\bar{\lambda} - \lambda||^{k+1}_{L_{k+1}(I)} \qquad (2.6.2)$$

where $g_k(\lambda,\bar{\lambda}) = \dfrac{1}{||\bar{\lambda}\lambda||^k_{L_\infty(I)}}$

Proof.

$$-\bar{\lambda}^k\lambda^k(\frac{1}{\lambda^k} - \frac{1}{\lambda^k})(\bar{\lambda} - \lambda) = (\bar{\lambda}^k - \lambda^k)(\bar{\lambda} - \lambda)$$

$$\geq (\tfrac{1}{2})^k\; |\bar{\lambda} - \lambda|^{k+1}$$

$$||\bar{\lambda}\lambda||^k_{L_\infty} \int_I - \left(\frac{1}{\bar{\lambda}^k} - \frac{1}{\lambda^k}\right)(\bar{\lambda} - \lambda) \geq \left(\frac{1}{2}\right)^k ||\bar{\lambda} - \lambda||^{k+1}_{L_{k+1}(I)}$$

Result follows.■

Theorem 2.5. Let stress σ be given by

$$\sigma(\lambda) = \sum_{k=M_1}^{M_2} \operatorname{sgn}(k) C_k \lambda^k \qquad M_1, M_2, k \in Z,\ C_k > 0 \tag{2.6.3}$$

Then

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), \bar{\lambda} - \lambda) \geq \sum_{k=M_1}^{M_2} \left(\frac{1}{2}\right)^{|k|} g_k(\bar{\lambda},\lambda) ||\bar{\lambda} - \lambda||^{|k|+1}_{L_{|k|+1}(I)}$$

where

$$g_k(\bar{\lambda},\lambda) = \left\{ \begin{array}{ll} C_k ||\bar{\lambda}\lambda||^k_{L_\infty(I)} & \text{for } k < 0 \\ 0 & \text{for } k = 0 \\ 1 & \text{for } k > 0 \end{array} \right] \tag{2.6.4}$$

Proof. The result is an immediate consequence of the definition of $\sigma(\lambda)$, Lemma 2.1 and Lemma 2.2.■

Theorem 2.6. Consider the stress $\sigma(\lambda)$ of (2.2.12) for type I problems for cases in which $C_k \geq 0$, $0 \leq k \leq p$. Then

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), \bar{\lambda} - \lambda) \geq \sum_{k=1}^{p} C_1 \left(\frac{1}{2}\right)^k ||\bar{\lambda} - \lambda||^{k+1}_{L_{k+1}(I)} \tag{2.6.5}$$

Proof. Result follows from Lemma 2.1.■

Corollary 2.4.

$$(\sigma(D\bar{u}) - \sigma(Du), D(\bar{u} - u)) \geq \sum_{k=1}^{p} C_k \left(\frac{1}{2}\right)^k ||\bar{u} - u||^{k+1}_{W^1_{k+1}(I)} \tag{2.6.6}$$

where $\bar{u}u \in \overset{\circ}{W}{}^1_{p+1}(I)$ and (,) denotes duality pairing on $L_{p+1}(I)$ and its dual.

where $\bar{u}$, $u \in \overset{\circ}{W}{}^1_{p+1}(I)$ and $(\cdot,\cdot)$ denotes duality pairing on $L_{p+1}(I)$ and its dual.■

We again cite as an example the Mooney material for problems of Type II.

Theorem 2.7. Let $W(\lambda)$ be given by the Mooney form with C_{10}, $C_{01} > 0$. Then the corresponding stress of (2.2.12) and (2.2.18) is strongly monotone on $L_2(I)$; i.e. there exists a constant γ_1 such that

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), \bar{\lambda} - \lambda) \geq \gamma_1 \| \bar{\lambda} - \lambda \|^2_{L_2(I)} \qquad (2.6.7)$$

Proof.

$$\sigma(\lambda) = -2C_{01}\lambda^{-3} - 2C_{10}\lambda^{-2} + 2C_{01} + 2C_{10}\lambda$$

We apply Lemma (2.2) and write

$$\begin{aligned}(\sigma(\bar{\lambda}) - \sigma(\lambda), \bar{\lambda} - \lambda) &\geq \left(\tfrac{1}{2}\right)^3 2C_{01}\, g_{-3}(\bar{\lambda},\lambda)||\bar{\lambda} - \lambda||^4_{L_4(I)} \\ &+ \left(\tfrac{1}{2}\right)^2 2C_{10}\, g_{-2}(\bar{\lambda},\lambda)||\bar{\lambda} - \lambda||^3_{L_3(I)} \\ &+ 2C_{10}||\bar{\lambda} - \lambda||^2_{L_2(I)}\end{aligned}$$

Therefore,

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), \bar{\lambda} - \lambda) \geq 2C_{10}||\bar{\lambda} - \lambda||^2_{L_2(I)}$$

Result follows by noting $y_1 = 2C_{10}$.■

Corollary 2.5. Let the conditions of Theorem 2.7 hold. Then

$$(\sigma(D\bar{u}) - \sigma(Du), D\bar{u} - Du) \geq \gamma_1 \|\bar{u} - u\|^2_{W_2^1(I)} \qquad \blacksquare$$

Theorem 2.8. Let $\bar{\lambda}, \lambda \in L_p(I)$, $p \geq 6$, with $0 < \bar{\lambda}, \lambda < \infty$. Let $\sigma(\lambda)$ be given by (2.2.12) with coefficients C_k, $-3 \leq k \leq 5$ corresponding to the Biderman form (2.2.19). Then, for certain C_k, there exists a positive constant β such that

$$(\sigma(\bar{\lambda}) - \sigma(\lambda), \bar{\lambda} - \lambda) \geq \beta \|\bar{\lambda} - \lambda\|^6_{L_6(I)} \tag{2.6.8}$$

wherein $(\cdot,\cdot)$ denotes duality pairing between $L_6(I)$ and $L_{6/5}(I)$.

Proof: The proof of this theorem follows the same lines as that of Theorem 2.7 and will, therefore, be omitted. $\blacksquare$

We again remark that λ can be replaced by $1 + Du$ and the above theorems can be written in terms of energy spaces $W_p^m(I)$ containing weak gradients of the displacements:

Corollary 2.6, Let the conditions of Theorem 2.7 hold. Then

$$(\sigma(D\bar{u}) - \sigma(Du), D(\bar{u} - u)) \geq \gamma \|\bar{u} - u\|^2_{W_2^1(I)} \tag{2.6.9}$$

where $\bar{u}, u \in \overset{\circ}{W}{}_2^1(I)$ and $(\cdot,\cdot)$ denotes the inner product on $L_2(I)$. $\blacksquare$

Corollary 2.7. Let the conditions of Theorem 3.7 hold. Then

$$(\sigma(D\bar{u}) - \sigma(Du), D(\bar{u} - u)) \geq \beta \|\bar{u} - u\|^6_{W_6^1(I)} \tag{2.6.10}$$

where $\bar{u}$, $u \in \overset{\circ}{W}{}^1_6(I)$ and $(\cdot,\cdot)$ denotes duality pairing on $L_6(I)$ and $L_{6/5}(I)$. ■

II. 7 Properties of σ for Very Regular Solutions. In the derivation of the mathematical properties of the stress σ in sections II.4 - II.6 we attempted to present the most general results possible. In these sections we adopted the following technique to obtain what we now call the most general results. We determined the minimal regularity of the exact solution u implied by the existence of some $W^1_p(I)$ norm of u. We attempted to derive continuity and strong monotonicity results in this $W^1_p(I)$ norm. If this was possible and if the norms of $\bar{u}$,u occuring in $g(\bar{u},u)$ (the "constant" in the continuity property) were in the $W^1_p(I)$ norm or some weaker norm, then we would say that these were the most general results possible. In every case except for the Mooney material (2.2.18) we achieved this objective. For the Mooney material consistent (meaning in the same norm) continuity and strongly monotone results occurred only in the $W^1_2(I)$ norm. Then the constant $g(\bar{u},u)$ contained much stronger norms of $\bar{u}$,u. Thus the Mooney material is a special case.

Suppose that the solution u, in addition to being in $W^1_p(I)$, is in some smaller space $W^1_q(I)$ for q>p. Then any theoretical work performed using the most general results will undoubtedly produce pessimistic conclusions. For this reason we would like to determine the effect on the continuity and monotonicity properties of σ of the process of increasing the regularity of the exact solution u in this sense.

In view of the discussions in II.2 and II.5, let us consider the following general form of constitutive relation.

$$\Sigma(u) = \sigma(\lambda) = \sum_{k=0}^{n} C_{-k}\lambda^{-k} + \sum_{j=1}^{p-1} C_j\lambda^j \tag{2.7.1}$$

where $\lambda = 1 + Du$, $n \geq 0$, $p \geq 2$.

We need to distinguish between positive and negative elastic constants C_j, $-n \leq j \leq p-1$. Toward this end, we denote

$$Z_n^+ = \{1 \leq k \leq n \quad , \quad C_{-k} \geq 0\}$$

$$Z_n^- = \{1 \leq k \leq n \quad , \quad C_{-k} < 0\}$$

$$Z_p^+ = \{1 \leq j \leq p-1, \quad C_j \geq 0\}$$

$$Z_p^- = \{1 \leq j \leq p-1, \quad C_j < 0\}$$

Either of these sets of integers can be empty, but they cannot be simultaneously empty. We now give a general theorem which goes in similar lines of one given by Oden and Reddy [64].

<u>Theorem 2.9</u>. Let α be the stress operator defined in (2.7.1) and let Z_n^+ be empty. Moreover, for α a positive real number, let

$$A(\alpha) = \sum_{j=-n}^{p-1} jC_j\alpha^{j-1} \tag{2.7.2}$$

Then, a necessary and sufficient condition that the stress operator be monotone is that

$$A(\alpha) \geq 0 \quad \forall \ \alpha \in \mathbb{R}^+ \tag{2.7.3}$$

In addition, let γ_r denote the number

$$\gamma_r = \inf_{\alpha \in \mathbb{R}^+} \begin{cases} A(\alpha) & , \quad r = 2 \\ \alpha^{2-r} B(\alpha) , & 2 < r \le p \end{cases} \tag{2.7.4}$$

where $B(\alpha)$ is defined as

$$B(\alpha) = \Big(\sum_{k \in Z_n^-} -C_{-k}\alpha^{-(k+1)} + \sum_{j \in Z_p^+} C_j \alpha^{j-1} + \sum_{j \in Z_p^-} jC_j \alpha^{j-1} \Big) \tag{2.7.5}$$

Then, if $\gamma_r > 0$

$$\langle \sigma(\lambda) - \sigma(\bar{\lambda}), \lambda - \lambda \rangle \ge \gamma_r \| \lambda - \bar{\lambda} \|^r_{L_r(I)} \tag{2.7.6}$$

Proof: By direct calculation, we find that the Gâteaux derivative of σ is continuous and is given by

$$\sigma'(\lambda) = A(\lambda)$$

where A is the function defined in (2.7.2). Then, by the theorem due to Minty [64], σ is monotone if (2.6.4) holds.

To establish (2.6.6), we directly use (2.6.2) to obtain

$$\langle \sigma(\lambda) - \sigma(\bar{\lambda}), \lambda - \bar{\lambda} \rangle = \int_I \Big(\sum_{k=1}^{n} - \frac{C_{-k}}{\lambda^k \bar{\lambda}^k} |\lambda^k - \bar{\lambda}^k| + \sum_{j=1}^{p-1} C_j |\lambda^j - \lambda^j| \Big) |\lambda - \bar{\lambda}| dx$$

$$= \int_I \Big(\sum_{k \in Z_n^-} - \frac{C_{-k}}{\lambda^k \bar{\lambda}^k} |\lambda^k - \bar{\lambda}^k| + \sum_{j \in Z_p^+} C_j |\lambda^j - \bar{\lambda}^j|$$

$$+ \sum_{j \in Z_p^-} C_j |\lambda^j - \bar{\lambda}^j| \Big) |\lambda - \bar{\lambda}| dx$$

where Z_n^+ is considered empty. Without loss of generality, let $\lambda > \bar{\lambda}$ (if $\lambda - \bar{\lambda}$, (2.6.6) is trivially satisfied). Further, let

$$\tau = \frac{\bar{\lambda}}{\lambda} < 1$$

and observe that

$$\langle\sigma(\lambda) - \sigma(\bar{\lambda}), \lambda - \bar{\lambda}\rangle = \int_I \Big(\sum_{k \in Z_n^-} - C_{-k}\tau^{-k}\bar{\lambda}^{-(k+1)} \frac{1-\tau^k}{1-\tau}$$

$$+ \sum_{j \in Z_p^+} C_j \; j\text{-}1 \; \frac{1-\tau^j}{1-\tau}$$

$$+ \sum_{j \in Z_p^-} C_j \lambda^{j-1} \frac{1-\tau^j}{1-\tau} \Big) |\lambda - \mu|^2 \, dx$$

for $r > 1$ we have $1 < (1-\tau^r)/(1-\tau) < r$ (for proof see [64]); therefore

$$\langle\sigma(\lambda) - \sigma(\bar{\lambda}), \lambda - \bar{\lambda}\rangle \geq \int_I B(\lambda)|\lambda - \bar{\lambda}|^2 \, dx$$

Next, we assume that $B(\lambda) > 0 \;\; \forall \lambda \in \mathbb{R}^+$, so that

$$\langle\sigma(\lambda) - \sigma(\bar{\lambda}), \lambda - \bar{\lambda}\rangle \geq \int_I \lambda^{2-r}(1-\tau)^{2-r} B(\lambda)|\lambda - \mu|^r \, dX$$

$$\geq \int_I \lambda^{2-r} B(\lambda)|\lambda - \mu|^r \, dX$$

for $2 < r \leq p$. For $r = 2$, we note that

$$\langle\sigma(\lambda) - \sigma(\bar{\lambda}), \lambda - \bar{\lambda}\rangle = \langle\sigma'(\hat{\lambda})\lambda - \bar{\lambda}, \lambda - \bar{\lambda}\rangle = \int_I A(\hat{\lambda})|\lambda - \bar{\lambda}|^2 \, dx$$

where $\hat{\lambda} = \theta\lambda + (1-\theta)\bar{\lambda} \quad \theta \in [0,1]$.

Now the function $f(\alpha) = \begin{cases} A(\alpha) & r = 2 \\ \alpha^{2-r}B(\alpha) & 2 < r \leq p \end{cases}$ is merely a function of the positive real numbers α and it may achieve a positive infimum independent of x. When an infimum exists, it is precisely the number γ_r defined on (2.6.5). Hence, the theorem follows. ■

Remark. 1. When Z_p^- is empty, we observe that application of Lemma 2.1 results in 2.7.6 with $r = p$, immediately.

2. When Z_n^+ is nonempty we get the infimum to be zero and hence strong monotonicity is not possible.

Theorem 2.10. Let σ be the stress operator defined in (2.7.1). Let $\lambda, \bar{\lambda}$ and μ be arbitrary elements in the space of admissible functions. Then

$$\langle \sigma(\lambda) - \sigma(\bar{\lambda}), \mu \rangle \leq E g_{st}(\lambda, \bar{\lambda}) ||\mu||_{L_s(I)} ||\lambda - \bar{\lambda}||_{L_t(I)} \tag{2.7.7}$$

wherein

$$E = \max_{-n \leq j \leq p-1} |C_j|; \quad 2 \leq s, \quad t \leq p \tag{2.7.8}$$

and

$$g_{st}(\lambda, \bar{\lambda}) = L^{\frac{1}{\alpha(s,t)}} + \sum_{k=1}^{n} ||(\lambda + \bar{\lambda})^{k-1}/\lambda^k \bar{\lambda}^k||_{L_{\alpha(s,t)}(I)} + \sum_{j=2}^{p-1} ||\lambda + \bar{\lambda}||^{j-1}_{L_{(j-1)\alpha(s,t)}(I)} \tag{2.7.9}$$

with $L = \text{mes}(I)$ and

$$\alpha(s,t) = st/(st - s - t) \tag{2.7.10}$$

Proof:

$$|\langle\sigma(\lambda)-\sigma(\bar\lambda),\mu\rangle| = \left|\int_I \Big(\sum_{k=1}^{n} C_{-k}(\lambda^{-k}-\bar\lambda^{-k}) + \sum_{j=1}^{p-1} C_j(\lambda^j-\bar\lambda^j)\Big)\mu\, dx\right|$$

$$\le \int_I \Big(\sum_{k=1}^{n} |C_{-k}||\lambda^{-k}-\bar\lambda^{-k}| + \sum_{j=1}^{p-1} |C_j||\lambda^j-\bar\lambda^j|\Big)\mu\, dx$$

$$\le E\int_I \Big(\sum_{k=1}^{n} \frac{|\lambda^k-\bar\lambda^k|}{\lambda^k\bar\lambda^k} + \sum_{j=1}^{p-1} |\lambda^j-\bar\lambda^j|\Big)\mu\, dx$$

$$\le E\int_I \Big(\sum_{k=1}^{n} \frac{(\lambda+\bar\lambda)^{k-1}}{\lambda^k\bar\lambda^k} + \sum_{j=1}^{p-1} (\lambda+\bar\lambda)^{j-1}\Big)|\lambda-\bar\lambda|\mu\, dx$$

$$\le E\Big|\Big|\sum_{k=1}^{n} \frac{(\lambda+\bar\lambda)^{k-1}}{\lambda^k\bar\lambda^k}$$

$$+ \sum_{j=1}^{p-1} (\lambda+\bar\lambda)^{j-1}\Big|\Big|_{L_{\alpha(s,t)}(I)} ||\mu||_{L_s(I)} ||\lambda-\bar\lambda||_{L_t(I)}$$

$$2 \le s\ , \quad t \le p$$

where $\frac{1}{\alpha(s,t)} + \frac{1}{s} + \frac{1}{t} = 1$ (or $\alpha(s,t) = st/(st-s-t)$).

Theorem follows by noting,

$$\Big|\Big|\sum_{k=1}^{n} \frac{(\lambda+\bar\lambda)^{k-1}}{\lambda^k\bar\lambda^k} + \sum_{j=1}^{p-1} (\lambda+\bar\lambda)^{j-1}\Big|\Big|_{L_{\alpha(s,t)}(I)} \le \sum_{k=1}^{n} \Big|\Big| \frac{(\lambda+\bar\lambda)^{k-1}}{\lambda^k\bar\lambda^k}\Big|\Big|_{L_{\alpha(s,t)}(I)}$$

$$+ L^{\frac{1}{\alpha(s,t)}} + \sum_{j=2}^{p-1} ||\lambda+\bar\lambda||^{j-1}_{L_{(j-1)\alpha(s,t)}(I)}$$

$$= g_{st}(\lambda,\bar\lambda)$$

Remark. Results for the Mooney, Biderman and IHT forms can be obtained by choice of n and p in (2.7.1) and proper substitution in Theorems 2.9 and 2.10.

CHAPTER III

APPROXIMATION OF THE ELASTOSTATIC PROBLEM FOR NONLINEAR ELASTIC SOLIDS

III.1 Introduction. In Chapter II, nonlinear elastostatic problems for compressible and incompressible elastic materials were defined. In this chapter we attempt to answer certain theoretical questions concerning the approximation of these problems. In particular we demonstrate convergence and determine the rate of convergence of the variational approximation (2.1.2) when implemented with finite element interpolants. In addition, we present the results of numerical experiments designed to verify these results.

In particular, we develop a priori error estimates for a number of important cases. In the first results which are for the case we have characterized in Chapter II as the most general results we obtain the same rate of convergence as that of linear problem for $p \geq 2$, and it is independent of the value of p. However, the theoretical rate obtained is $h^{\frac{k}{p-1}}$. For linear interpolation this can be improved to h^{ν} where $\nu = k + \frac{1}{p} - \frac{1}{2}$ [64]. In particular, when p = 2, corresponding to the linear theory, we obtain an error of order h^k, which agrees with recent results in linear approximation theory (e.g. [60]).

Secondly, for problems in which the exact solution is significantly more regular than the exact solution for the most general case, we find the same result. When the exact solution $u \in W^1_\infty(I)$, the rate of con-

vergence for the nonlinear problem is the same as the one for the linear problem. That is, the approximation error in the $W_2^1(I)$ norm is of order h^k where k is the degree of the polynomials in the finite element basis. *The rate of convergence is independent of the power* p of the displacement gradient appearing in the strain energy function.

We comment on the reason we put so much emphasis on the approximation of the elastostatic problem here. This question naturally arises since this work is primarily devoted to the elastodynamic problem and, in particular, wave propagation. The reason is that the theoretical results for the dynamic problems are based on the static problem and in particular on the static error estimates. Thus the static problem is a necessary but interesting diversion.

III.2 The Finite-Element Approximation. We are now ready to consider Galerkin approximations of boundary value problems of Types I and II in which the approximation subspaces are endowed with the interpolation properties of the finite element method.

In particular, we consider nonlinear elasticity problems of the following variational form;

$$(\sigma(Du), Dv) = \ell(v), \quad \forall\, v \in \overset{\circ}{W}{}_p^1(I) \tag{3.2.1}$$

where, for conditions set in Theorem 2.5 and Theorem 2.6, for some $p > 1$,

$$\langle \sigma(Du) - \sigma(Dv), w\rangle \leq G_{st}(u,v)\,||u - v||_{W_t^1(I)}\,||w||_{W_s^1(I)} \tag{3.2.2}$$

$$\langle \sigma(Du) - \sigma(Dv), D(u - v)\rangle \geq \gamma\,||u - v||^p_{W_p^1(I)} \tag{3.2.3}$$

where γ is a constant > 0, G(u,v) is a nonnegative function of gradients of u and v (e.g. $G(u,v) = Eg(1 + Du, 1 + Dv)$ where $g(\cdot,\cdot)$, s and t are described in Theorem 2.6. In (3.2.1), $\ell(v)$ is a continuous linear functional on $\overset{\circ}{W}{}^1_p(I)$ defined by

$$\ell(v) = \langle f,v\rangle \ , \quad v \in \overset{\circ}{W}{}^1_p(I) \tag{3.2.4}$$

where the data f is the body force component introduced in (2.1.1).

Recall that I is the set of particles $\{X: \ X \in \mathbb{R}, \ 0 \le X \le a < \infty\}$. Our Galerkin approximation of (3.2.1) involves the following construction: we construct a partition P of I defined by the set of nodes $\{X^\alpha\}_{\alpha=0}^G$ such that

$$0 = X^0 < X^1 < X^2 < \ . \ . \ . < X^G = a$$

We denote

$$h_\alpha = X^\alpha - X^{\alpha-1}$$

$$h = \max_{1\le\alpha\le G} \{h_\alpha\}$$

and assume that the mesh is quasi-uniform; i.e., there exists a real number $r \ge 1$ such that $r \ge h/(\min_\alpha\{h_\alpha\})$. The closed subintervals $\overline{I}_\alpha = \{X: \ X^{\alpha-1} \le X \le X^\alpha, \ 1 \le \alpha \le G\}$ are interpreted as finite elements and

$$I = \bigcup_{\alpha=1}^{G} \overline{I}_\alpha, \quad I_\alpha \bigcap I_\beta = \phi, \ \alpha \ne \beta$$

Over each element I_α we construct local interpolation functions $\psi_N^{(\alpha)}(X)$, N = 1,2, using the techniques described in [62]. Upon connect-

ing the elements, we obtain a system of global interpolation functions $\{\phi_\alpha(X)\}_{\alpha=1}^{G}$ which form the basis of a G-dimensional subspace of $\overset{\circ}{W}{}_p^1(I)$ denoted $S_h^k(I)$. The functions $\phi_\alpha(X)$ are piecewise polynomials of some degree $k \geq 1$ and are continuous on I with derivatives of order $m \geq 1$ in $L_p(I)$. We shall assume, therefore, that $S_h^k(I)$ is endowed with the fundamental finite-element interpolation properties [62,63].

(i) Let $P_k(I)$ denote the space of polynomials of degree $\leq k$ on I. Then the finite element basis functions $\{\phi_\alpha(X)\}_{\alpha=1}^{G}$ define a linear mapping Π_h from $W_p^{k+1}(I)$ onto $S_h^k(I)$ such that

$$\Pi_h v = v \text{ if } v \in P_k(I) \tag{3.2.5}$$

(ii) If v is an arbitrary element in $W_p^{k+1}(I)$

$$\inf_{V \in S_h^k(I)} \| v - V \|_{W_p^m(I)} \leq C_1 h^{k+1-m} |v|_{W_p^{k+1}(I)} \tag{3.2.6}$$

where C_1 is a positive constant independent of v and h, $m \geq 1$, and $|v|_{W_p^{k+1}(I)}$ denotes the semi-norm

$$|v|_{W_p^{k+1}(I)} = \left\{ \int_I \sum_{i=k+1} |D^i v|^p \, dX \right\}^{1/p} \tag{3.2.7}$$

It follows that every function $V(X)$ in $S_h(I)$ is of the form

$$V(X) = \sum_{\alpha=1}^{G} B^{\alpha}\phi_{\alpha}(X)$$

where $\{B^{\alpha}\} \in \mathbb{R}^G$. The finite-element-Galerkin approximation U of the solution u of (3.2.1) is the unique function

$$U(X) = \sum_{\alpha=1}^{G} A^{\alpha}\phi_{\alpha}(X) \in S_h^k(I) \subset W_p^1(I) \qquad (3.2.8)$$

such that

$$(\sigma(DU), DV) = \ell(V), \quad \forall\ V \in S_h^k(I) \qquad (3.2.9)$$

We also observe from (3.2.1) and (3.2.10) the property of the orthogonality of error.

$$(\sigma(Du) - \sigma(DU), DV) = 0 \qquad \forall\ V \in S_h^k(I) \qquad (3.2.10)$$

where u is the exact solution and U its finite element approximation. Obviously, the introduction of (2.2.8) into (2.2.9) leads to a system of nonlinear algebraic equations in the coefficients A^{α}:

$$(\sigma(D\sum_{\alpha} A^{\alpha}\phi_{\alpha}), D\phi_{\beta}) = \ell(\phi_{\beta}), \quad 1 \leq \alpha,\beta < G \qquad (3.2.11)$$

For stress operators of the form (2.2.12), we remark that

$$(\sigma(D\sum_{\alpha} A^{\alpha}\phi_{\alpha}), D\phi_{\beta}) = \sum_{k=M_1}^{M_2} c_k \int_I (1 + \sum_{\alpha} A^{\alpha}D\phi_{\alpha})^k D\phi_{\beta}dX$$

$$1 \leq \alpha,\ \beta \leq G \qquad (3.2.12)$$

Techniques for solving (3.2.11) numerically are discussed in [61].

All of these preliminaries bring us to the fundamental approximation theorem :

Theorem 3.1. Let the stress operator $\sigma(D(\cdot))$ satisfy (3.2.2) and (3.2.3) for some $p \geq 1$, for conditions stated in Theorems 2.9 and 2.10. Let u, the exact solution of nonlinear variational elasticity problem (3.2.1), belong to $W_p^\ell(I)$ and let U denote its finite element approximation; i.e. U is the element of the space $S_h^k(I) \subset W_p^1(I)$ with properties (3.2.6) - (3.2.8). Then there exists a constant $C > 0$ independent of the mesh parameter h, such that

$$||e||_{W_r^1(I)} \leq C(h^\mu + h^{\mu_1} H(u))||u||_{W_p^\ell(I)} = O(h^\nu) \tag{3.2.13}$$

where $2 \leq r \leq p$ and e is the approximation error

$$e = u - U \tag{3.2.14}$$

$$\mu = \min(k, \ell - 1) \tag{3.2.15}$$

$$\nu = \min(\mu, \mu_1),\ \mu_1 = \frac{\mu}{r-1} \text{ and } H(u) = \lim_{h \to 0} \left(\frac{\tilde{C}_1 G_{rr}(u, \tilde{U})}{\gamma} ||u||^{2-r}_{W_p^\ell(I)} \right)^{1/r-1}$$

$\tilde{U} \in S_h^k(I)$ is the projection of u, satisfying (3.2.7).

Proof:

$$||e||_{W_r^1(I)} \leq ||u - \tilde{U}||_{W_r^1(I)} + ||U - \tilde{U}||_{W_r^1(I)}$$

$$\leq \hat{C}_1 h^\mu ||u||_{W_p^\ell(I)} + ||U - \tilde{U}||_{W_r^1(I)}$$

From (3.2.3), we can write

$$||U - \tilde{U}||^r_{W^1_r(I)} \le \frac{1}{\gamma} (\sigma(DU) - \sigma(D\tilde{U}),D(U - \tilde{U}))$$

$$= \frac{1}{\gamma} (\sigma(Du) - \sigma(D\tilde{U}),D(U - \tilde{U}))$$

$$\le \frac{1}{\gamma} G_{rr}(u,\tilde{U})||u - \tilde{U}||_{W^1_r(I)} ||U - \tilde{U}||_{W^1_r(I)}$$

We have used the property (3.2.11) in the above derivation. Substituting above,

$$||U - U||_{W^1_r(I)} \le [\frac{\tilde{C}_1 G_{rr}(u,\tilde{U})}{\gamma} ||u||_{W_p(I)}]^{\frac{1}{r-1}} h^{\frac{\mu}{r-1}}$$

$$G_{rr}(u,\tilde{U}) = G_{rr}(u,u - (u - \tilde{U}))$$

We denote $H(u) = \lim_{h\to 0} [\frac{\tilde{C}_1 G_{rr}(u,\tilde{U})}{\gamma}]^{\frac{1}{r-1}} ||u||^{\frac{2-r}{r-1}}_{W^\ell_p(I)}$

The result follows immediately with $\mu_1 = \frac{\mu}{r-1}$ and $\nu = \min(\mu,\mu_1)$.∎

We observe that since $p > 1$, the rate-of-convergence of the finite element approximation will be governed by the second term on the right side of the inequality (3.2.13); i.e., for approximations employing piecewise polynomials of degree k,

$$||e||_{W^1_p(I)} = O(h^\nu)$$

Oden and Reddy have shown that for piecewise linear interpolation ($\mu = 1$), $\nu = \frac{1}{2} + \frac{1}{r}$, $r \ge 2$, which is clearly a sharper rate-of-convergence.

Also for $r = 2$, we observe that the convergence rate is optimal in $||\cdot||_{W_2^1(I)}$.

III.3 The Nonlinear Nitsche Trick. The estimate (3.2.13) describes the behavior of the approximation error in the $W_p^1(I)$-norm; i.e., it represents an error in an energy-type norm or an L_p-estimate of the error in the stress. It is natural to now inquire how the error in displacements behaves. The analogous question for linear problems was elusive for a number of years until Nitsche [65] proposed a technique, using regularity theory, for obtaining L_2 error estimates. Similar results were obtained by Aubin [66]. Nitsche's procedure has become known as the "Nitsche trick" (see, e.g. [67]). We shall now demonstrate an extension of Nitsche's idea to nonlinear problems of the type (3.2.1), which might be termed a "nonlinear Nitsche trick" [64].

We note that if u, v, and w are arbitrary elements of the space of admissible displacements, the stress operator (2.7.1) obeys the Lagrange formula

$$\langle \Sigma(u) - \Sigma(v), w \rangle = \langle D\Sigma(\theta u + (1 - \theta)v) \cdot (u - v), w \rangle \qquad (3.3.1)$$

where $\theta \in (0,1)$.

Our extended Aubin-Nitsche procedure involves the analysis of a linear variational boundary-value problem on Banach spaces. For this purpose, we will need to call upon the following theorem, the proof of which can be found in Nečas [68].

Lemma 3.1. Let U and V be two real reflexive Banach spaces and $B: U \times V \to \mathbb{R}$ a bilinear form such that

$$|B(u,v)| \leq M ||u||_U ||v||_V \quad \forall\, u \in U, \quad \forall\, v \in V \tag{3.3.2}$$

$$\inf_{||u||_U = 1} \ \sup_{||v||_V \leq 1} |B(u,v)| \geq \gamma > 0$$

$$\sup_{u \in U} |B(u,v)| > 0 \quad v \neq 0 \tag{3.3.4}$$

where M and γ are positive constants. Then there exists a unique element $u_0 \in U$ such that

$$B(u_0, v) = f(v) \quad \forall\, v \in V \tag{3.3.5}$$

where $f \in V'$; i.e., f is a continuous linear functional on V. Moreover,

$$||u_0||_U \leq \frac{1}{\gamma} ||f||_{V'} \tag{3.3.6}$$

We next apply Lemma 8.1 to obtain some special results on existence and regularity of solutions to a differential equation defined on Banach spaces.

Lemma 3.2. Let $a \in L_\infty(I)$ and $a(x) \geq K > 0$ almost everywhere in I, where $I \subset \mathbb{R}$ and K is a constant > 0. Let $\psi \in L_q(I)$, $q = p/(p-1)$, $\infty > p \geq 2$. Then there exists a unique function $w_0 \in W_q^2(I) \cap \mathring{W}_q^1(I)$ such that

$$\langle a w_0', v' \rangle_0 = \langle \psi, v \rangle_0 \quad \forall\, v \in \mathring{W}_p^1(I) \tag{3.3.7}$$

where $\langle \cdot,\cdot \rangle_0$ *denotes duality pairing on* $L_q(I) \times L_p(I)$.

Moreover, there exists a positive constant $\tilde{C}$ such that:

$$||w_o||_{W_q^2(I)} \leq \tilde{C}||\psi||_{L_q(I)} \tag{3.3.8}$$

Proof : Clearly, from Holder's inequality, $|\langle aw',v'\rangle| \leq A_o||Dw||_{L_q(I)} \times ||Dv||_{L_p(I)} \leq A_o||w||_{W_q^1(I)}||v||_{W_p^1(I)}$ where $A_o = \underset{x\in I}{\text{ess sup}}\,|a|$. Thus, the form $B(w,v) \equiv \langle aw',v'\rangle$ is a continuous bilinear form from $\overset{\circ}{W}{}_q^1(I) \times \overset{\circ}{W}{}_p^1(I)$ into $\mathbb{R}$, i.e. (3.3.2) is satisfied with $M = A_o$.

Now observe that $|\langle aw',v'\rangle| \geq K\langle w',v'\rangle \;\; \forall\, v \in \overset{\circ}{W}{}_p^1(I)$ and suppose $\infty > p > 2$. Then $p \geq q = q/(p-1)$. Consequently, $W_p^1(I) \subseteq W_q^1(I)$. Thus, for a given $w \in \overset{\circ}{W}{}_q^1(I)$, pick a special $v = \hat{v} \in \overset{\circ}{W}{}_p^1(I)$ such that almost everywhere in I, $\hat{v}' = |w'|^{q-2}\,w'$. Then $|\hat{v}'|^p = |w'|^q \in L_1(I)$, so this choice can be made. Consequently,

$$|B(w,\hat{v})| = |\langle aw',\hat{v}'\rangle| \geq K||w'||^q_{L_q(I)} = K||w||^q_{W_q^1(I)}$$

Thus

$$|B(w, \frac{\hat{v}}{||\hat{v}||_{W_p^1(I)}})| \geq K\frac{||w||^q_{W_q^1(I)}}{||w||^{q/p}_{W_q^1(I)}} = K||w||_{W_q^1(I)}$$

Therefore,

$$\inf_{||w||_{W_q^1(I)} = 1}\;\sup_{||v||_{W_p^1(I)} \leq 1} |B(w,v)| \geq K > 0$$

and the form in (3.3.7) satisfies (3.3.3).

To establish (3.3.4), pick a special $w = \hat{w}$ such that $\hat{w}' = |v'|^{p-2}\,v'$ which is certainly possible. Then

$$\sup_{w \in W_q^1(I)} |B(w,v)| \geq |B(\hat{w},v)| \geq K||v||^p_{W_p^1(I)} > 0 \qquad v \neq 0$$

so that (3.3.4) is also satisfied. Consequently, a unique solution w_o to (3.3.7) exists by virtue of Lemma 3.1.

To show that $w_o \in W_q^2(I)$, pick a test function ϕ in $\mathcal{D}(I)$ (recall that the space $\mathcal{D}(I)$ of test functions is dense in $\overset{\circ}{W}{}_p^1(I)$. Then

$$\langle aw_o', \phi' \rangle = \langle \psi, \phi \rangle \qquad \forall \phi \in$$

where now $aw_o' \in \mathcal{D}(I)'$; i.e. aw_o' is a distribution. Thus $\langle (aw_o')' - \psi, \phi \rangle = 0 \;\; \forall \phi \in \mathcal{D}(I)$, and since $\psi \in L_q(I)$, so does $aw_o'' + a'w_o'$ where these derivatives are interpreted in the sense of distributions. Thus w_o has distributional derivatives of order ≤ 2 in $L_q(I)$ and this is equivalent to saying that $w_o \in W_q^2(I)$ for any $\psi \in L_q(I)$.

Now, in view of (3.3.6), $||w_o||_{W_q^1(I)} \leq K^{-1}||\psi||_{L_q(I)}$. Since

$$||w_o||^q_{W_q^2(I)} = ||w_o''||^q_{L_q(I)} + ||w_o||^q_{W_q^1(I)} \leq ||w_o''||^q_{L_q(I)} + K^{-q}||\psi||^q_{L_q(I)}$$

and

$$\begin{aligned} ||w_o''||_{L_q(I)} &= ||\tfrac{1}{a}\psi - \tfrac{a'}{a}\, w_o'||_{L_q(I)} \\ &\leq C_o||\psi||_{L_q(I)} + C_1||w_o'||_{L_q(I)} \\ &\leq C_o||\psi||_{L_q(I)} + C_1||w_o||_{W_q^1(I)} \\ &\leq (C_o + C_1K^{-1})||\psi||_{L_q(I)} \end{aligned}$$

we have

$$||w_o||^q_{W^2_q(I)} \le [(C_o + C_1K^{-1})^q + K^{-q}]||\psi||^q_{L_q(I)} = \tilde{C}^q||\psi||^q_{L_q(I)}$$

Whence (3.3.8). ■

With these preliminaries behind us, we can now establish the following theorem :

Theorem 3.2. Let the following conditions hold :

(i) the stress operator $\Sigma(u)$ (2.7.1) be coercive, hemicontinuous, and strictly monotone.

(ii) the exact solution u of 2.1.2 be in $W^{\ell}_p(I)$ with $\ell > 2 + 1/p$.

(iii) the finite element approximation U of u exists, is unique, and is in the space $S^k_h(I)$, and estimate (3.2.13) holds. Then the approximation error e = u - U satisfies, for sufficiently small h,

$$||e||_{L_p(I)} \le \hat{C}h^{\nu+1}||u||_{W_p(I)} G_{qp}(u,U) \tag{3.3.9}$$

where $\hat{C}$ is a positive constant independent of h and μ and $G_{qp}(u,U)$ are defined in (3.2.15) and (3.2.2), respectively.

Proof : We begin by considering the auxiliary linear variational problem with variable coefficients

$$\langle D\Sigma(\bar{u})w',v'\rangle = \langle \psi,v\rangle \quad \forall\, v \in \overset{\circ}{W}{}^1_p(I) \tag{3.3.10}$$

where $\psi \in L_q(I)$ ($q = p/p-1$)) and

$$\bar{u} = \theta u + (1-\theta)U = U + \theta e\ , \quad \theta \in (0,1)$$

By hopothesis, $\Sigma(\bar{u}) > \beta > 0$ a.e. in I. And, by the Sobolev imbedding

theorem, u is sufficiently regular that $\Sigma(\bar{u}) \in L_\infty(I)$. Hence, by Lemma 3.2, there exists a unique solution w_o to (3.3.10) in $W_q^2(I) \cap \overset{\circ}{W}{}_q^1(I)$ such that

$$||w_o||_{W_q^2(I)} \le \tilde{C}||\psi||_{L_q(I)}$$

Now, let us define

$$\psi = \frac{|e|^{p-2}\, e}{||e||^{p-1}_{L_p(I)}}$$

Then,

$$\langle \psi, e \rangle = ||e||^{1-p}_{L_p(I)}\, ||e||^{p}_{L_p(I)} = ||e||_{L_p(I)}$$

and

$$||\psi||_{L_q(I)} = 1$$

Thus, setting v = e in (3.3.10) and using the Lagrange formula (3.3.1), we have

$$\begin{aligned} ||e||_{L_p(I)} &= \langle D\Sigma(\bar{u}) \cdot w_o', e' \rangle_o \\ &= \langle D\Sigma(\bar{u}) \cdot e', w_o' \rangle_o \\ &= \langle \Sigma(u) - \Sigma(U), w_o' \rangle_o \\ &= \langle \Sigma(u) = \Sigma(U), w_o' - \tilde{W}_o' \rangle_o \end{aligned}$$

where, in the last step, we have used the orthogonality condition and $\tilde{W}_o$ is the projection of w_o in $S_h^k(I)$ and, therefore, satisfies

$$||w_o - \tilde{W}_o||_{W_q^1(I)} \le Ch||w_o||_{W_q^2(I)} \le C\tilde{C}h||\psi||_{L_q(I)} = C_1 h \qquad (3.3.11)$$

Hence, using (3.2.2), we have

$$||e||_{L_p(I)} \leq ||u - U||_{W_p^1(I)} ||w_o - \tilde{W}_o||_{W_q^1(I)} G_{qp}(u,U)$$

$$\leq Ch^{\nu} ||u||_{W_p^{\ell}(I)} C_1 h G_{qp}(u,U)$$

which gives (3.3.9). ■

We observe that the estimate (3.3.9) is not an a priori estimate since the bound is a function of U. However, we can state an obvious result for cases when u is sufficiently smooth.

Corollary 3.1. Let the conditions of Theorem 3.1 hold and let

$$\lim_{h \to 0} G_{qp}(u,U) = H(u) \qquad (3.3.12)$$

Then

$$||e||_{L_p(I)} \leq \hat{C} h^{\nu+1} ||u||_{W_p^{\ell}(I)} \tilde{H}(u) \qquad (3.3.13)$$

III.4 Convergence and Accuracy for Solutions with Increased Regularity.

If the exact solution to (3.2.1) has increased regularity over the regularity assumed in section III.3 then the results of III.3 are quite pessimistic. In this section we develop error estimates for one such case of increased regularity. In particular we consider the solution of (3.2.1) for a Type I material. In this case the Piola-Kirchhoff stress operator is of the general form (2.2.12) with $M_1 = 0$ and $M_2 = p-1$. Normally the minimal regularity of the solution in this case is $u \in W_p^1(I)$. We will assume here however, that $u \in W_\infty^1(I)$. Of course, this is a rather special case and different results would be obtained

for other regularity hypothesis. However, we note that at this point we do not know how to do this in general. It is possible that the methods of Tarter [69] are applicable.

In this case we have continuity and strong monitonicity properties for (3.2.1) of the following form:

$$\langle \sigma(Du) - \sigma(Dv), w \rangle \leq G_{p2}(u,v) \|u - v\|_{W_p^1(I)} \|w\|_{W_2^1(I)} \tag{3.4.1}$$

$$(\sigma(Du) - \sigma(Dv),\ Du-Dv) \geq \gamma \ \|u-v\|^2_{W_2^1(I)} \tag{3.4.2}$$

where γ is a positive constant and $G_{p2}(u,v)$ is a nonnegative function of the gradients of u and v (e.g. $G(u,v) = g_{p2}(1+Du,1+Dv)$ where $g_{p2}(\cdot,\cdot)$ is given in Theorem 2.10.

We can then state the fundamental approximation result for Type I materials in the case when the solution has increased regularity (and in particular $u \in W^1_\infty(I)$):

Theorem 3.3. Let u be the solution of the nonlinear variational elasticity problem (3.2.1) and let $u \in W^1_\infty(I)$. In addition, let U denote the finite element approximation to u; i.e., U is the element of the space $S_h^k(I) \in W_p^1(I)$ which satisfies (3.2.5) and (3.2.6) that satisfies (3.2.9), wherein the stress operator σ satisfies (3.4.1) and (3.4.2). Then there exists a positive constant C independent of the mesh parameter h, such that

$$\|e\|_{W_2^1(I)} \leq C[1 + \frac{H(u)}{\gamma}]\ h^k \|u\|_{W_2^{k+1}(I)} \tag{3.4.3}$$

where e is the approximation error

$$e = u - U$$

Proof. The theorem follows immediately from Theorem 3.1.■

The next result extends the nonlinear Nitsche trick to the case in which the solution has increased regularity.

Theorem 3.4. Let the conditions of Theorem 3.3 hold. In particular, let $u \in W^1_\infty(I)$. In addition let the stress operator $\sigma(D(\cdot))$ be Gateaux differentiable on a convex subset Ω of $W^1_p(I)$ containing the solution u to (3.2.1) and its finite element approximation U. Then the approximation error e is such that

$$||e||_{L_2(I)} \leq \hat{C}h^{k+1}||u||_{W_2^{k+1}(I)}^{\tilde{H}(u)} \tag{3.4.4}$$

Proof. The result follows immediately from theorem 3.2.■

III.5 Numerical Experiments. The theoretical results of sections III.2, III.3, and III.4 indicate the performance of variational methods in nonlinear elasticity problems with various parameters of the problem. In this section we attempt to verify the theoretical results experimentally.

We consider here the simplest nonlinear elasticity problem for the Type I material. We let the stress σ have the following form:

$$\sigma(u_x) = C_0 + C_1(1+u_x) + C_2(1+u_x)^2 \tag{3.5.1}$$

Then we solve the boundary value problem

$$- \frac{d}{dX}\sigma = f$$

$$u(0) = 0 \tag{3.5.2}$$

$$u(1) = 0$$

We approximate (3.5.1) with (3.2.11), and we use piecewise linear basis functions (k=1). We use a constant mesh size h. In terms of h the basis functions take the following form:

$$\phi_1(X) = 1 - \frac{X}{h} \qquad 0 \leq X \leq h$$

$$\phi_\alpha(X) = \begin{cases} \frac{X}{h} - (\alpha-2) & (\alpha-2)h \leq X \leq (\alpha-1)h \\ \alpha - \frac{X}{h} & (\alpha-1)h \leq X \leq \alpha h \end{cases} \tag{3.5.3}$$

$$\alpha=2, \ldots, N-1$$

$$\phi_N(X) = \frac{X}{h} - (N-2) \qquad (N-2)h \leq X \leq (N-1)h$$

where h is the discretization parameter and

$$h = \frac{1}{(N-1)}$$

In order to solve (3.2.11) we introduce two iterative schemes:

(i) Operator Splitting. Using (3.5.1) in conjunction with (3.2.10) we find that

$$\int_0^L (C_1 + 2C_2 + C_2 \sum_\gamma A_\gamma A\phi_\gamma) \sum_\alpha A_\alpha D\phi_\alpha D\phi_\beta dx = f_\beta \tag{3.5.4}$$

where $f_\beta = \int_0^L f\phi_\beta dX$. We create an iterative scheme by evaluating the quantity in parenthesis on the left-hand side in (3.5.4) at iteration point n and the other quantity on the left-hand side in (3.5.4) at iteration point n+1. Then we find that

$$\int_0^L \left(C_1 + 2C_2 + C_2 \sum_\gamma A^n D\phi_\gamma\right) \sum_\alpha A_\alpha^{n+1} D\phi_\alpha D\phi_\beta dX = f_\beta \qquad (3.5.5)$$

In matrix form the system is

$$[T^n]\ \{A^{n+1}\} = \{f\} \qquad (3.5.6)$$

where

$$[T^n] = \begin{bmatrix} k_1^n+k_2^n & -k_2^n & & & & & & \\ -k_2^n & k_2^n+k_3^n & -k_3^n & & & & & \\ & -k_3^n & k_3^n+k_4^n & -k_4^n & & & & \\ & & \cdot & \cdot & & & & \\ & & & \cdot & \cdot & & & \\ & & & & \cdot & \cdot & & \\ & & & & -k_{N-3}^n & k_{N-3}^n+k_{N-2}^n & -k_{N-2}^n & \\ & & & & & -k_{N-2}^n & k_{N-2}^n+k_{N-1}^n & -k_{N-1}^n \\ & & & & & & -k_{N-1}^n & k_{N-2}^n+k_{N-1}^n \end{bmatrix}$$

and

$$k_i^n = K_o - K_1 A_i^n + K_1 A_{i+1}^n \tag{3.5.7}$$

where

$$K_o = \frac{C_1 + 2C_2}{h}$$

$$K_1 = \frac{C_2}{h^2}$$

This iteration could be started from an initial guess. However, we have had good results by setting $A_i^o = 0$, $1 \leq i \leq N$.

Now let U be the approximate solution and U^n be the approximate solution at the nth iteration. We need to show that as $n \to \infty$, $U^n \to U$. (3.5.4) and (3.5.5) are equivalent to

$$[(C_1 + 2C_2 + C_2 DU)DU, DV] = (f,V) \quad \forall\ V \in S_h(\Omega) \tag{3.5.8}$$

and

$$[(C_1 + 2C_2 + C_2 DU^n)DU^{n+1}, DV] = (f,V) \quad \forall\ V \in S_h(\Omega) \tag{3.5.9}$$

where $[\cdot,\cdot]$ denotes the L_2 inner-product. Now we let $e_h^j = U - U^j$. Then subtracting (3.5.9) from (3.5.8) and letting $V = e_h^{n+1}$

$$[C_2 e_h^n DU, De_h^{n+1}] + [(C_1 + 2C_2 + C_2 DU^n)De_h^{n+1}, De_h^{n+1}] = 0 \tag{3.5.10}$$

Using continuity and coercivity properties for the linear problem we have constructed, we get

$$\| e_h^{n+1} \|_{H^1(0,1)} \leq \eta^n \| e_h^n \|_{H^1(0,1)} \tag{3.5.11}$$

where

$$\eta^n = \frac{C_2 \|DU\|_{L_\infty(0,1)}}{\inf\limits_{X \in (0,1)} (C_1+2C_2DU^n)}$$

Clearly the iterative scheme is contractive if $DU \in L_\infty(0,1)$, $|DU^n| < C_1+2C_2/C_2$, and $\eta^n < 1$ for all n. These are rather restrictive conditions and we have obtained good results for much more general problems with this sceme.

(ii) Newton Raphson Method. The system of nonlinear algebraic equations can be written in the form

$$\underset{\sim}{F}(\underset{\sim}{A}) = 0 \tag{3.5.12}$$

where

$$F = \begin{Bmatrix} g_1^2 + g_2^1 - f_2 \\ g_2^2 + g_3^1 - f_3 \\ \vdots \\ g_{N-1}^2 - g_N^1 - f_N \end{Bmatrix}$$

and

$$g_i^1 = K_oA_i - K_oA_{i+1} - K_1A_i^2 - K_1A_{i+1}^2 \tag{3.5.13}$$

$$g_i^2 = -g_i^1$$

Clearly the Jacobian matrix $\underset{\sim}{J} = [\partial F_i/\partial A_j]$ is of the form

$$\underset{\sim}{J} = \begin{bmatrix} J_1+J_2 & -J_2 & & & & & \\ -J_2 & J_2+J_3 & -J_3 & & & & \\ & -J_3 & J_3+J_4 & -J_4 & & & \\ & & \cdot & \cdot & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ & & & & -J_{N-3} & J_{N-3}J_{N-2} & -J_{N-2} \\ & & & & & -J_{N-2} & J_{N-2}+J_{N-1} & -J_{N-1} \\ & & & & & & -J_{N-1} & J_{N-1}+J_n \end{bmatrix}$$

where

$$J_i = K_o - 2K_1A_i + 2K_1A_{i+1}$$

Then we use a series expansion to construct the iterative scheme

$$\underset{\sim}{A}^{n+1} = \underset{\sim}{A}^n - (\underset{\sim}{J}^n)^{-1}\underset{\sim}{F}^n \tag{3.5.14}$$

The initial numerical tests were designed to verify the results of section III.4. In particular we attempt to verify the rate of convergence for solutions with increased regularity. We choose as the exact solution the function $u = x - x^3$ on the interval [0,1]. Clearly, $u \in W^1_\infty(0,1) \cap W^2_2(0,1)$. Thus the results of Theorem 3.3 and Theorem 3.5 apply. Using these results we know that if $u \in W^1_\infty(0,1) \cap W^2_2(0,1)$, then the errors in the energy and displacement are given as follows:

$$\|e\|_{W_2^1(0,1)} \leq O(h)$$
$$\|e\|_{L_2(0,1)} \leq O(h^2) \qquad (3.5.15)$$

In Figure 3.1 we show the exact solution compared to the approximate solution using four finite elements. In Figure 3.2 we present the experimental convergence and rate of convergence of the approximation error e in the $W_2^1(0,1)$ norm. In Figure 3.3 we present the experimental convergence and rate of convergence of the approximation error e in the $L_2(0,1)$ norm. The experimental results clearly support and are even identical to the theoretical predictions of section III.4.

In the second set of numerical experiments we attempted to verify the <u>most general</u> estimates of sections III.2 and III.3. In this case we need to choose as the exact solution a function u which is $W_3^1(0,1)$ but is not in $W_4^1(0,1)$. We choose in this case the function $u = x^{-3/4} - x$. We can easily show that $u \in W_3^1(0,1)$ and $u \notin W_4^1(0,1)$. Thus this solution satisfies the constraint $u \notin W_\infty^1(0,1)$ and we can proceed.

We note initially that due to the stress singularity at the origin $u \notin W_3^2(0,1)$, but $u \in W_4^{4/3}(0,1)$. This is an intermediate space between $W_3^1(0,1)$ and $W_3^2(0,1)$. In this case the error of the nonlinear approximation is

$$\|e\|_{W_3^1(0,1)} \leq O(h^{1/3}) \qquad (3.5.16)$$

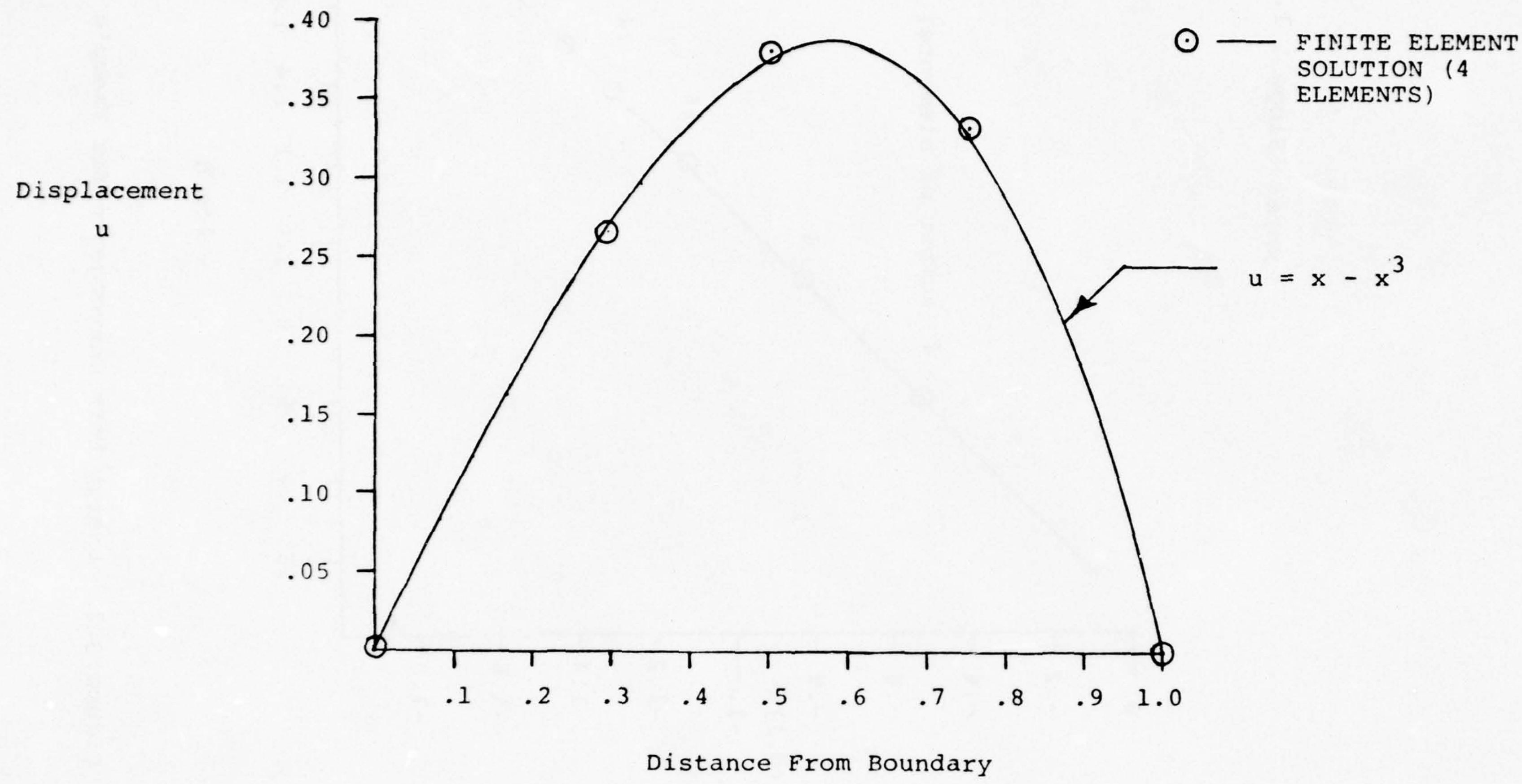


Figure 3.1. Exact solution for Example 1 compared with a four element approximation.

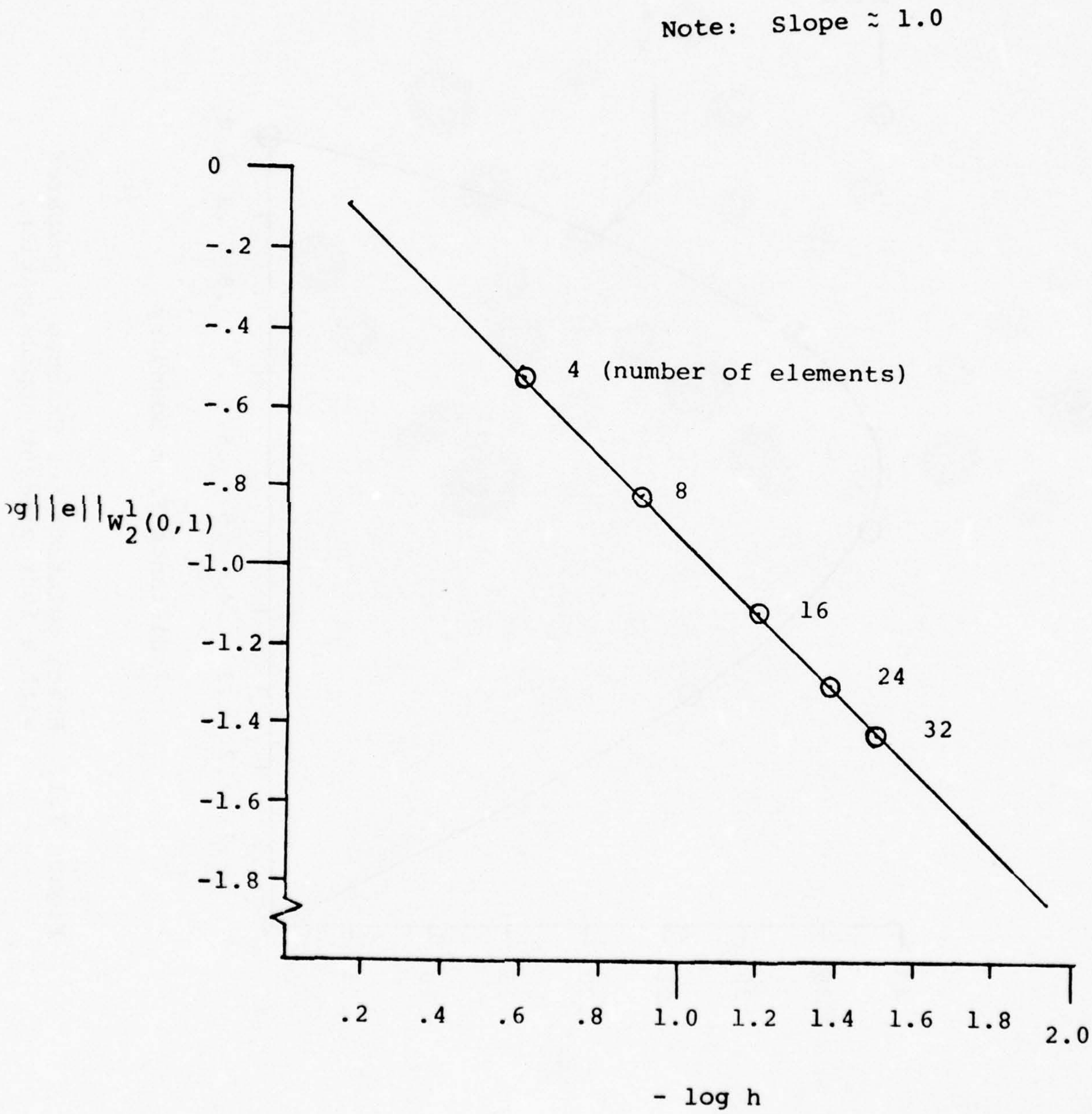


Figure 3.2. Energy Norm Convergence for Example 1.

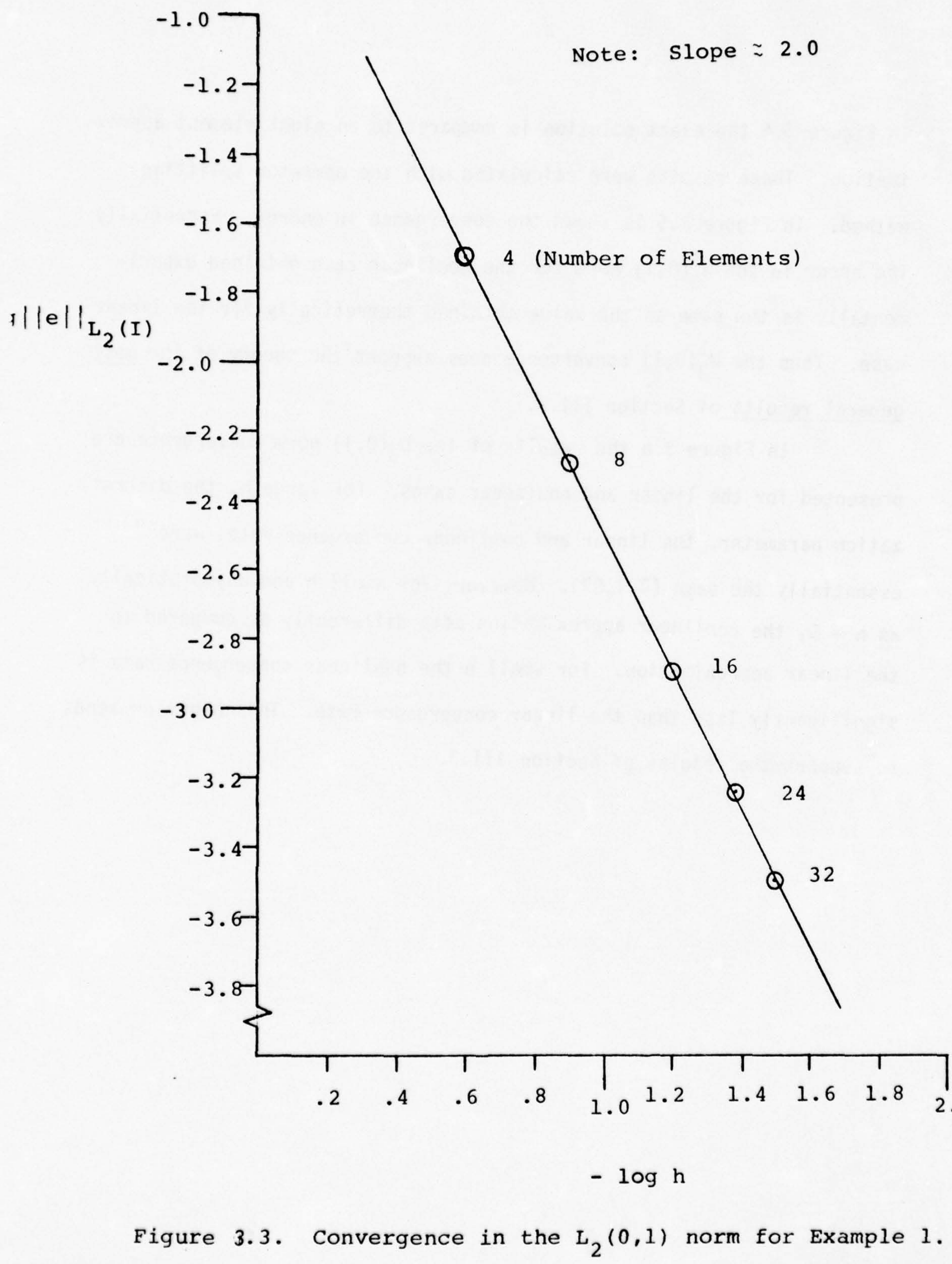


Figure 3.3. Convergence in the $L_2(0,1)$ norm for Example 1.

In Figure 3.4 the exact solution is compared to an eight element approximation. These results were calculated with the operator splitting method. In Figure 3.5 is shown the convergence in energy. Essentially the error in the $W_3^1(0,1)$ norm for the nonlinear case obtained experimentally is the same as the value obtained theoretically for the linear case. Thus the $W_3^1(0,1)$ convergence does support the theory of the most general results of Section III.2.

In Figure 3.6 the results of the $L_3(0,1)$ norm convergence are presented for the linear and nonlinear cases. For large h, the discretization parameter, the linear and nonlinear convergence rates were essentially the same ($\simeq$ 1.07). However, for small h and asymptotically as $h \to 0$, the nonlinear approximation acts differently as compared to the linear approximation. For small h the nonlinear convergence rate is significantly less than the linear convergence rate. This behavior tends to support the results of Section III.2.

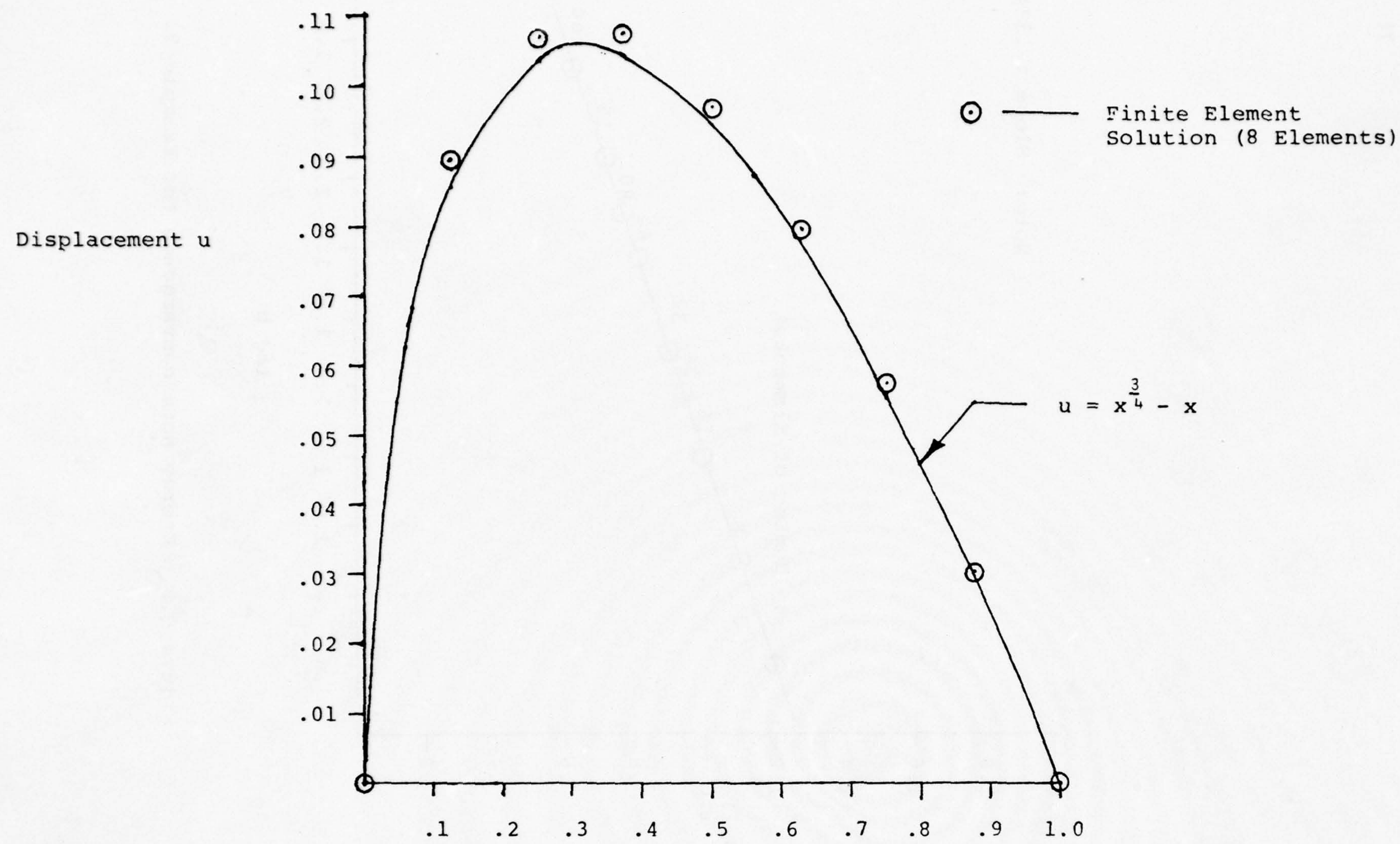


Figure 3.4. Exact solution for Example 2 compared with an eight element approximation.

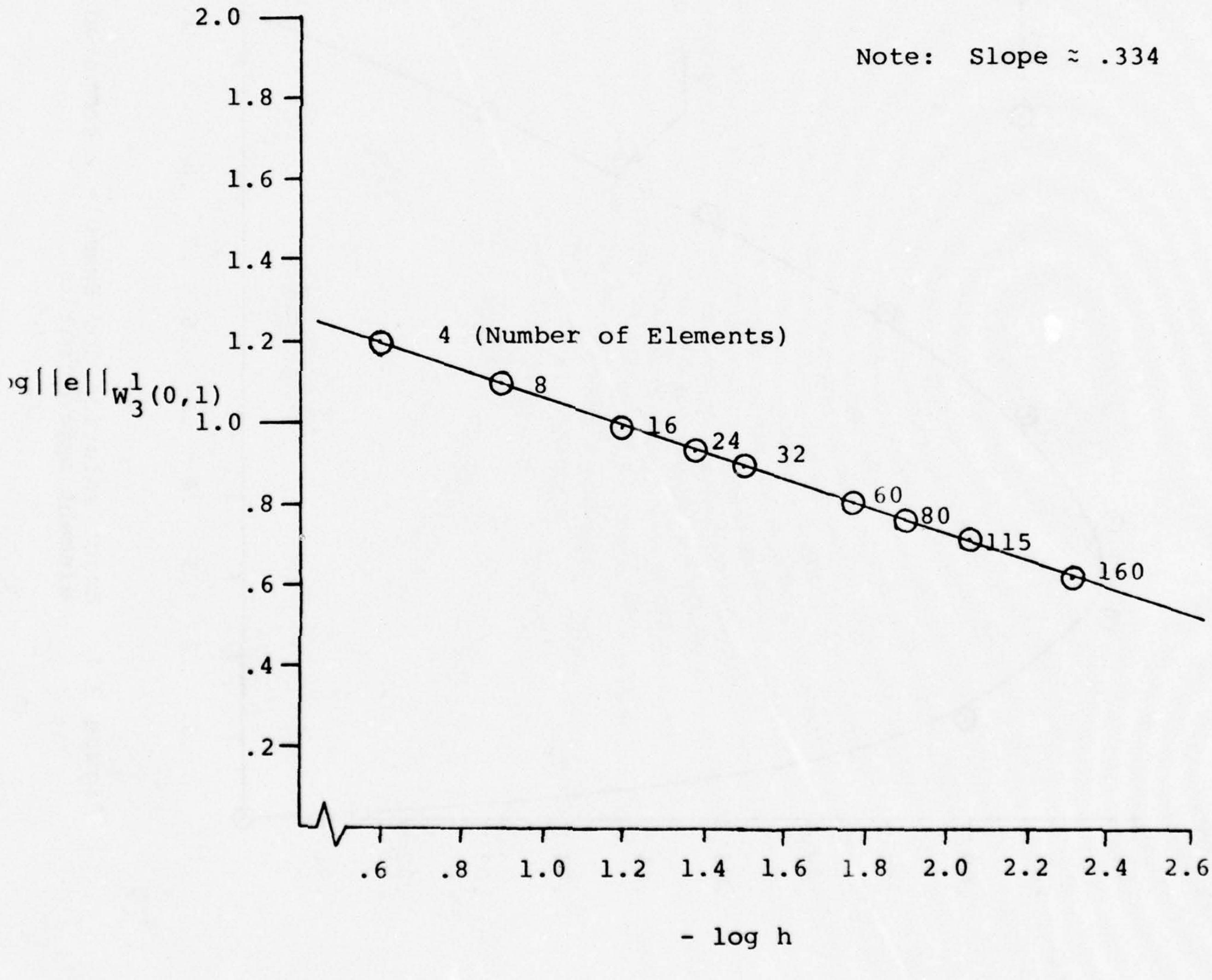


Figure 3.5. Energy norm convergence for Example 2.

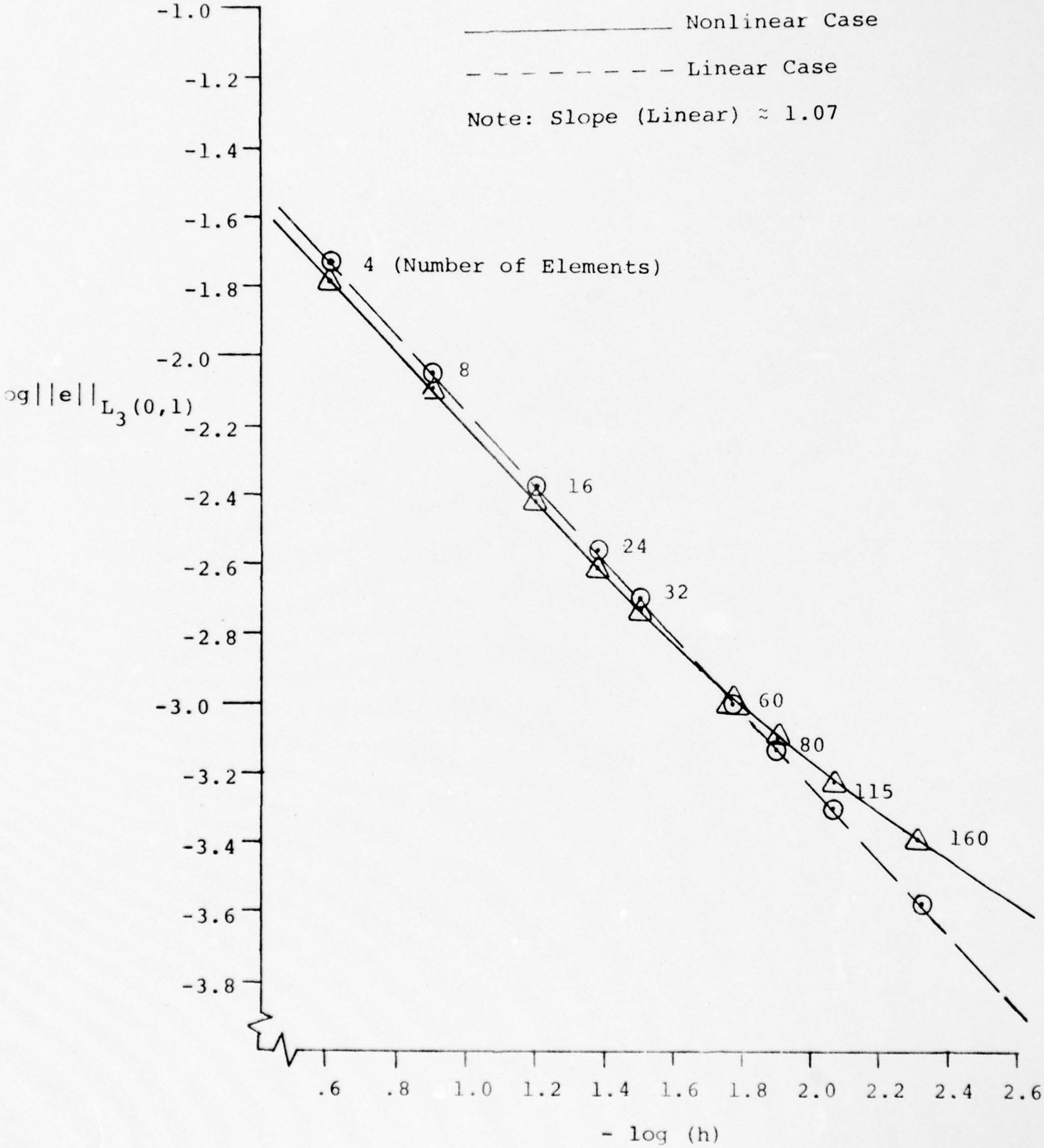


Figure 3.6. Convergence in the $L_3(I)$ norm for Example 2.

CHAPTER IV

A THEORY OF GENERALIZED GALERKIN METHODS OF THE SHOCK FITTING TYPE

IV.1 Introduction. In this section we develop a theory of generalized Galerkin models. These models represent the extension of the classical Galerkin methods to free boundary problems and use discontinuous trial functions. Of course, the physical application is to shock wave problems. We show that generalized Galerkin models in conjunction with discontinuous finite element interpolants lead to shock fitting schemes. We develop accuracy estimates, demonstrate convergence, and develop stability criteria for both semidiscrete and fully discretized shock fitting schemes.

IV.2 Some Additional Notation. In this chapter, we use some special notation not introduced in our preliminary remarks on the subject in Section II.3. In particular, we are now familiar with the Sobolev space $W_p^m(I)$ of functions whose generalized derivatives of order less than or equal to m or in $L_p(I)$ in the spatial variables. However, temporal behavior of a function u(X,t) may also be in a space $L_q(0,T)$; thus we write

$$\|u\|_{L_q(0,T;W_p^m(I))} = \left\{ \int_0^T \|u(\cdot,t)\|^q_{W_p^m(I)} \, dt \right\}^{1/q} \qquad (4.2.1)$$

Thus, $L_2(0,T;H^m(I))$ denotes the space of bivariate functions measurable and square-integrable in the temporal variable over $(0,T) \subset \mathbb{R}$ and, for each $t \in (0,T)$, are in $H^m(I)$. Likewise, we use the notation

$$\|u\|_{L_\infty(0,T;W_q^m(I))} = \underset{0<t<T}{\text{ess. sup.}} \|u(\cdot,t)\|_{W_q^m(I)} \tag{4.2.2}$$

Since we shall usually be concerned with solutions over some specific time interval $(0,T)$, we shall sometimes omit the interval designation and simply write $L_p(W_q^m(I))$ for $L_p(0,T;W_q^m(I))$.

Now suppose Z is an arbitrary point of I and K, $L \subset I$ are open sets bounding the point Z. Then, as special notations, we let $(\cdot,\cdot)_Z$ denote the scalar product of two function evaluated at Z and we define a "boundary norm" associated with Z by

$$|||u|||_{Z(K,L)} = \sup_{X \in K \cup L} |u(X)| \tag{4.2.3}$$

In addition, we let $((\cdot,\cdot))_{L_2(I)}$ denote the L_2 inner product.

Other notation will be defined where it first appears in subsequent articles.

IV.3 Mechanical Formulation. We now consider construction of the elastodynamics problem of Class I or II for the case in which N - 1 shock waves may exist at any time $t \geq 0$ in the material. We continue to denote by I a possibly unbounded set of material particles equivalent to an open subset of $\mathbb{R}$. Let $\{Y_k\}_{k=0}^{N} \equiv Q$ denote a set of N + 1 real valued functions from $[0,T] \to \mathbb{R}$ into $\mathbb{R}$ such that for each $t \in [0,T]$, Q is a partition of I; i.e., if $0 = \inf\{I\}$, $L_o = \sup\{I\} \leq \infty$, then

$$0 = Y_0 < Y_1 < Y_2 \ldots < Y_N = L_o$$

At $t \in [0,T]$, we denote by $J_i(t)$ the open sets

$$J_i(t) = (Y_{i-1}(t),\ Y_i(t)),\quad 1 \leq i \leq N \tag{4.3.1}$$

and

$$I/Q = \bigcup_{i=1}^{N} J_i \qquad \forall\ t \in [0,T]$$

In particular Q subdivides I into a number of disjoint open sets of particles J_i at each $t \in [0,T]$. These sets shall correspond to <u>shockless</u> <u>subdomains</u> of B at time t, and the functions $Y_i(t)$, which together with $\sum_i J_i$ describe the closure of $\sum_i J_i$, represent particles at which there may exist <u>shocks</u>--i.e., surfaces of discontinuity in the displacement gradients $u_X(X,t)$ and the velocity $\dot{u}(X,t)$.

With these conventions in mind, the basic physical conservation laws take on the following form:

(i) Balance of Linear Momentum.

$$\sum_{i=1}^{N} \int_{J_i} [\ \frac{\partial}{\partial t}(\rho\dot{u}) - \rho f]dX + \sum_{i=1}^{N-1} \rho\ V_i[\![\dot{u}]\!]_{Y_i} = \sigma(L_o,t) - \sigma(0,t) \tag{4.3.2}$$

(ii) Conservation of Energy.

$$1/2 \sum_{i=1}^{N} \int_{J_i} [\frac{\partial}{\partial t}(\rho\dot{u}^2 + 2e) - 2\rho f\dot{u}]dX + 1/2 \sum_{i=1}^{N-1} \rho V_i[\![\dot{u}^2]\!]_{Y_i}$$

$$+ \sum_{i=1}^{N-1} V_i[\![e]\!]_{Y_i} = \{\sigma(X,t)\dot{u}(X,t) + q(X,t)\}\Big|_{X=0}^{X=L_o} \tag{4.3.3}$$

(iii) The Clausius-Duhen Inequality

$$\sum_{i=1}^{N} \int_{J_i} \frac{\partial \xi}{\partial t}\, dX + \sum_{i=1}^{N-1} V_i [\![\xi]\!]_{Y_i} \geq \left\{\frac{q}{\theta}\right\} \Bigg|_{X=0}^{X=L_o} \tag{4.3.4}$$

Here V_i is the intrinsic speed of the i^{th} wave,

$$V_i(t) = \frac{dY_i(t)}{dt} \tag{4.3.5}$$

$[\![\psi]\!]_{Y_i}$ denotes the jump in any field quantity at the surface Y_i,

$$[\![\psi]\!]_{Y_i} = \psi(Y_i^-,t) - \psi(Y_i^+,t) \tag{4.3.6}$$

$u = u(X,t)$, $\sigma(X,t)$, $\hat{u}(X,t)$ are the displacement, stress and velocity at X at time t. f is the body force per unit mass, e is the internal energy per unit of reference volume, $q(X,t)$ is the heat flux, ξ the entropy, and is the absolute temperature. We also use the obvious notation,

$$\{\psi(x,t)\}\Bigg|_{X=0}^{X=L_o} = \psi(L_o,t) - \psi(0,t) \tag{4.3.7}$$

In <u>classical</u> formulations of wave problems in mechanics, it is argued that within each open set J_i the integrands in (4.3.2), (4.3.3) and (4.3.4) are sufficiently smooth to give meaning to a <u>pointwise statement of the wave problem</u>; that is, together with certain boundary- and initial-conditions, it is asserted that the set $\{u,\sigma,e,q,\xi,\theta\}$ is such that

$$\left.\begin{aligned} \rho\ddot{u} - \sigma_X &= \rho f \\ \dot{e} - \sigma u_X - q_X &= 0 \\ \theta\xi - q_X + \frac{q\theta_X}{\theta} &\geq 0 \end{aligned}\right\} \quad \begin{aligned} &\forall \ (X,t) \in J_i(t) \times (0,T] \\ &1 \leq i \leq N \end{aligned} \tag{4.3.8}$$

$$\left.\begin{aligned} \rho V_i [\![\dot{u}]\!]_{Y_i} + [\![\sigma]\!]_{Y_i} &= 0 \\ 1/2\, \rho V_i [\![\dot{u}^2]\!]_{Y_i} + V_i [\![e]\!]_{Y_i} + [\![\sigma\dot{u}]\!]_{Y_i} + [\![q]\!]_{Y_i} &= 0 \\ V_i [\![\xi]\!]_{Y_i} - [\![\tfrac{q}{\theta}]\!]_{Y_i} &\geq 0 \end{aligned}\right\} \quad \begin{aligned} &\text{on } Y_i(t) \times (0,T], \\ &1 \leq i \leq N-1 \end{aligned} \tag{4.3.9}$$

<u>IV.4 Galerkin Methods for Waves with Multiple Shocks</u>. We can introduce an alternate form for the global energy balance (4.3.3). This form is the basis for a variational theory of shock propagation and natural Galerkin approximation associated with the variational theory.

Initially we introduce the global energy balance for a shockless region J_i in the form

$$\frac{d}{dt}\frac{1}{2}\int_{J_i} (\rho\dot{u}^2 + 2e)dX = \{\sigma\dot{u} + q\}\Big|_{X = Y_{i-1}(t)}^{X = Y_i(t)} + \int_{J_i} \rho f\dot{u}dX \tag{4.4.1}$$

where the notation $\{\cdot\}\Big|_{X=Y_{i-1}(t)}^{X=Y_i(t)}$ means the difference in the average value of the quantity in parenthesis on the singular surfaces Y_i and Y_{i-1}. Differentiating in the first term in (4.4.1), introducing $(4.3.8)_2$ and reducing the term involving q_X to a boundary integral, we get

$$\int_{J_i} (\rho\ddot{u}\dot{u} + \sigma\dot{u}_X)dX + V_i\{\tfrac{1}{2}\rho\dot{u}^2 + e\}\Big|_{X=Y_{i-1}(t^+)}^{X=Y_i(t^-)}$$

$$= \{\sigma\dot{u}\}\Big|_{X=Y_{i-1}(t)}^{X=Y_i(t)} - \Delta q_i + \Delta q_{i-1} + \int_{J_i} \rho f\dot{u}dX$$

Here

$$\Delta q_i = \{q\}\Big|_{X=Y_i(t)}^{X=Y_i(t^-)} \tag{4.4.2}$$

Thus, we may now sum to obtain for the entire domain I,

$$\sum_{i=1}^{N} \int_{J_i} (\rho\ddot{u}\dot{u} + \sigma\dot{u}_X)dX + \sum_{i=1}^{N-1} V_i(\tfrac{1}{2}\rho [\![\dot{u}^2]\!]_{Y_i} + [\![e]\!]_{Y_i})$$

$$= \sigma\dot{u}\Big|_{X=0}^{X=L_o} - \sum_{i=1}^{N-1} ([\![q]\!]_{Y_i} - \int_{J_i} \rho f\dot{u}dX) \tag{4.4.3}$$

Now, for any field $\psi(X,t)$ let $\overline{\psi}_i$ denote the <u>average</u> <u>value</u> at the surface Y_i:

$$\overline{\psi}_i = \frac{1}{2}(\psi(Y_i^+,t) + \psi(Y_i^-,t)) \tag{4.4.4}$$

Let $[\![\dot{u}]\!]_{Y_i} \neq 0$, then

$$\frac{1}{2}\rho V_i [\![\dot{u}^2]\!]_{Y_i} = -\frac{1}{2}[\![\sigma]\!]_{Y_i} \frac{[\![\dot{u}^2]\!]_{Y_i}}{[\![\dot{u}]\!]_{Y_i}}$$

$$= -[\![\sigma]\!]_{Y_i}\,\overline{\dot{u}}_i \tag{4.4.5}$$

Moreover, since

$$-[\![\sigma]\!]_{Y_i}\,\overline{\dot{u}}_i + [\![\sigma\dot{u}]\!]_{Y_i} = \overline{\sigma}_i\,[\![\dot{u}]\!]_{Y_i} \tag{4.4.6}$$

$(4.3.9)_2$ can be written as

$$V_i[\![e]\!]_{Y_i} + \overline{\sigma}\,[\![\dot{u}]\!]_{Y_i} + [\![q]\!]_{Y_i} = 0 \tag{4.4.7}$$

Now let us assume that (4.4.7) is satisfied. Then e and q and their jumps can be eliminated from (4.4.3) to give the final global equation,

$$\sum_{i=1}^{N} \int_{J_i} (\rho\ddot{u}\hat{u} + \sigma\hat{u}_X)dX + \sum_{i=1}^{N-1} (\tfrac{1}{2}\rho V_i [\![\dot{u}^2]\!]_{Y_i} - \bar{\sigma}_i [\![\hat{u}]\!]_{Y_i})$$

$$= \{\sigma\hat{u}\}\Big|_{X=0}^{X=L_0} + \sum_{i=1}^{N} \int_{J_i} \rho f\hat{u}dX \qquad (4.4.8)$$

This equation is a <u>weak conservation form</u> of the condition of balance of linear mementum in the material body with jump terms. This result is our primary tool in developing Galerkin methods. The central question that arises is the following: if the <u>weak conservation form</u> (4.4.9) is used for problems of shock propagation, are the jump conditions (4.3.9) satisfied at the shock surfaces $Y_i(t)$, $1 \leq i \leq N-1$? The form (4.4.8) is developed from the global energy balance (4.3.3). Thus, the local energy jump condition $(4.3.9)_2$ is certainly satisfied (at least in a weak or average sense) using the form (4.5.8) because $(4.3.9)_2$ is developed from a global energy balance by shrinking the volume to zero at the shock surface. In addition, in the derivation of (4.5.8) it was assumed that the alternate jump condition (4.5.7) was identically satisfied on the surfaces $Y_i(t)$, $1 \leq i \leq N-1$. But since (4.5.7) is satisfied exactly and $(4.3.9)_2$ is satisfied in an average sense, the local momentum jump condition $(4.3.9)_1$ is satisfied in an average (see the derivation of (4.5.7)). In addition, the jump condition corresponding to the

Clausius-Duhem inequality $(4.3.9)_2$ is intrinsically satisfied because of the constitutive assumptions introduced in Section *II.2*.

Another from of (4.5.8) can be obtained by integrating by parts in the first term on the lefthand side, using (4.5.6) and the identity

$$\frac{1}{2}[\![\dot{u}^2]\!]_{\gamma_i} = [\![\dot{u}]\!]_{\gamma_i}\,\dot{\bar{u}}_i$$

We get

$$\sum_{i=1}^{N}\int_{J_i}(\rho\ddot{u} - \sigma_X - \rho f)\dot{u}\,dx + \sum_{i=1}^{N-1}(\rho V_i[\![\dot{u}]\!]_{\gamma_i} + [\![\sigma]\!]_{\gamma_i})\dot{\bar{u}}_i = 0 \qquad (4.4.9)$$

Now we denote by M_1 the space containing velocities $\dot{u}$ which satisfy the kinematical constraints of the problem under consideration. Since $\dot{u}$ is arbitrary in (4.4.9)

$$\sum_{i=1}^{N}\int_{J_i}(\rho\ddot{u} - \sigma_X - \rho f)v\,dX + \sum_{i=1}^{N-1}(\rho V_i[\![\dot{u}]\!]_{\gamma_i} + [\![\sigma]\!]_{\gamma_i})\bar{v}_i = 0,$$

$$\forall\ v \in M_1 \qquad (4.4.10)$$

The first term in (4.4.10) is the weak version of the momentum equation $(4.3.8)_1$ applied over a union of domains which do not contain singular surfaces internally. The second term represents a weak version

(in the same spirit as the first term) of the momentum jump condition $(4.3.9)_1$. However, the weighting function is the average value at the shock surface. Note that it is an implicit part of the this formulation that the weight functions v be taken from the space of velocities which are, in general, discontinuous at the Y_i, $1 \leq i \leq N-1$. This point marks the digression from the classical weak formulation of the partial differential equations involved.

In the classical weak formulation we choose as the space of test functions M_2 which are in general at least continuously differentiable everywhere in I. Then the weak form of the problem would be to find that u(X,t) such that

$$\int_I (\rho \ddot{u} - \sigma_X - \rho f) v dX = 0 \quad , \quad \forall \; v \in M_2 \tag{4.4.11}$$

We believe that Galerkin procedures constructed from (4.4.11) portray too smooth a solution of the problem to be useful for shock wave calculations. That is, many of the most useful and economical Galerkin schemes constructed from (4.4.11) do not converge to shock wave solutions. On the other hand, we maintain (and subsequently prove) that a Galerkin procedure consistent with (4.4.10) can converge to shock wave solutions.

In order to construct Galerkin approximations consistent with (4.4.10) and (4.4.11), we introduce further notation. In addition to the partition Q described earlier, let P denote a partition of I defined by the set of material nodal points $\{X_i\}_{i=0}^{M}$ where

$$0 = X_0 < X_1 < X_2 < \ldots < X_M = L_o$$

Let

$$h_i = X_i - X_{i-1} \quad , \quad 1 \leq i \leq M$$

and

$$I_i = [X_{i-1}, X_i] = \{X : X \in I,\ X_{i-1} \leq X \leq X_i\}$$

so that

$$I = \bigcup_{i=1}^{M} I_i$$

We assume that P is <u>quasi-uniform</u>, i.e., there is a constant $\gamma > 0$ such that

$$\gamma \leq h_i \leq h \quad , \quad 1 \leq i \leq M$$

where h is the <u>mesh parameter</u>

$$h = \sup_{1 \leq i \leq M} \{h_i\}$$

We also denote by $P_k(I)$ the space of polynomials of degree $\leq k$ on I.

We next introduce two finite-element-Galerkin subspaces of $W_p^m(I)$:

$$M_k^m(I,P) = \{v \mid v \in C^m(I) \cap P_k(I_i),\ 1 \le i \le M\}$$

$$H_k^m(I,P,Q) = \{v \mid v \in L_2(I) \cap C^m(J_j) \cap P_k(q_{ij});$$

$$1 \le j \le N,\ 1 \le i \le M,\ q_{ij} = I_i \cap J_j\}$$

Two Galerkin schemes can now be defined.

I. Shock Fitting Scheme. Find that $U \in \mathring{H}_k^m(I,P,Q) \cap C^0(I) \subset M_1$ such that

$$\sum_{i=1}^{N} ((\rho\ddot{U},W))_{L_2(J_i)} + \sum_{i=1}^{N} ((\sigma,W_X))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N-1} [\![\sigma W]\!]_{Y_i} + \sum_{i=1}^{N-1} (\rho V\,[\![\dot{u}]\!] + [\![\sigma]\!]\ ,\bar{W})_{Y_i}$$

$$= \sum_{i=1}^{N} ((\rho f,W))_{L_2(J_i)} \quad \forall\ W \in H_k^m(I,P,Q) \tag{4.4.12}$$

II. Shock Smearing Scheme. Find that $U \in M_k^m(I,P) \subset M_2$ such that

$$((\rho\ddot{U},W))_{L_2(I)} + ((\sigma,W_X))_{L_2(I)} = ((\rho f,W))_{L_2(I)}, \quad \forall\ W \in M_k^m(I,P) \tag{4.4.13}$$

In the above equations, the initial conditions are defined through either an L_2 or H^1 projection into the appropriate subspace. It should be understood that the term "shock smearing" in II is used only to indicate that the approximation is endowed with a smoothness which is not characteristic of the exact solution to the problem.

A temporal approximation can be developed by introducing a partition R of the interval [0,T] defined by $R = \{t_0, t_1, \ldots, t_r\}$ where $0 = t_0 < t_1 < \ldots < t_r = T$ and $t_{n+1} - t_n = \Delta t$, $n = 0, \ldots, r-1$. Then the sequence $\{U^n\}_{n=0}^{r}$ represents the values of the function $U(t)$ $\mathring{H}_k^m(I,P,Q)$ evaluated at the points of the partition R. The central difference operator is defined by

$$\delta_t^2 U^n = \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} \tag{4.4.14}$$

We assume that the Galerkin approximation satisfies the kinematical compatibility equation of the first order

$$[\![\dot{U}]\!]_{\gamma_i} = V_i [\![U_X]\!]_{\gamma_i}$$

Using this relationship, (4.4.12) and (4.4.14), we obtain a <u>fully discretized shock fitting scheme</u>:

$$\sum_{i=1}^{N} ((\rho\delta_t^2 \, U^n, W))_{L_2(J_i)} + \sum_{i=1}^{N} ((\sigma(U_X^n), W_X))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N-1} [\![\sigma(U_X^n) \, W]\!]_{Y_i} + \sum_{i=1}^{N-1} (-p(V_i^n)^2 [\![U_X^n]\!] + [\![\sigma(U_X^n)]\!] \;, \bar{W})_{Y_i}$$

$$+ \sum_{i=1}^{N} ((\rho f, W))_{L_2(J_i)} = 0 \qquad \forall \quad W \in \mathring{H}_k^m(I,P,Q) \qquad (4.4.15)$$

IV.5 Some Fundamental Approximation Theory Lemmas. Initially we define a function Z in a subspace $S_h(I)$ M, the original solution space, through the nonlinear "energy projection."

$$((\sigma(u_X) - \sigma(Z_Y) \;, V_X))_{L_2(I)} = 0 \quad \forall \quad V \in S_h(I) \qquad (4.5.1)$$

In later developments we choose to identify $S_h(I)$ with either $H_k^m(I,P,Q)$ or $M_k^m(I,P)$. We require that $S_h(I)$ satisfy a finite element interpolation property

$$\inf_{V \in S_h^k(I)} ||u-V||_{W_{p+1}^{\ell}(I)} \leq C \, h^{j-\ell} \, ||u||_{W_{p+1}^{j}(I)} \qquad \ell \leq j \leq k+1$$

Then a fundamental approximation theory problem is to determine the magnitude of the so called "interpolation error" E = u - Z and its temporal derivatives in various norms. Estimates of this kind for the Type I materials of Chapter III will be presented in a series of Lemmas generally termed approximation theory results.

AD-A031 238   TEXAS UNIV AT AUSTIN DEPT OF AEROSPACE ENGINEERING AN--ETC   F/G 20/11
A STUDY OF CONVERGENCE AND STABILITY OF FINITE ELEMENT APPROXIM--ETC(U)
AUG 76   J T ODEN, L C WELLFORD, C T REDDY   DAAG29-76-G-0022

UNCLASSIFIED   ARO-11860.9-M   NL

2 OF 4
ADA031238

If we notice that (4.5.1) is a nonlinear elliptic boundary value problem in weak form we can use the results of Chapter III to arrive at the first two Lemmas which represent estimates for E in $W_p^1(I)$ and $L_p(I)$.

Lemma 4.1. Suppose Z is the element in $S_h^k(I) \subset W_r^1(I)$, $2 \leq r \leq p$, which satisfies (4.5.1) and $\sigma(\cdot)$ satisfies Theorem 2.9 and 2.10. Then there exists a positive constant C_1 independent of the mesh parameter h, such that

$$||E||_{W_r^1(I)} \leq C_1(h^{\mu} + h^{\hat{\mu}} H(u))||u||_{W_p^{\ell}(I)} \tag{4.5.2a}$$

where $\mu = \min(k,\ell-1)$ and $H(\underset{\sim}{u})$ is defined in Theorem 3.1, $\hat{\mu} = \frac{\mu}{r-1}$.

For piecewise linear approximation we can show that $\hat{\mu}$ can be improved to $\frac{1}{2} + \frac{1}{r}$ [64], for $\mu = 1$. Denoting $\nu = \min(\mu,\mu_1)$,

$$||E||_{W_r^1(I)} \leq C(u)h^{\nu}||u||_{W_p^{\ell}(I)} \tag{4.5.2b}$$

where C(u) is a constant independent of h.

Lemma 4.2. Let the conditions of Lemma 4.1 hold, and let $\sigma(\cdot)$ be Gâteaux differentiable on a convex subset Ω of $W_r^1(I)$ containing u and Z. Then there exists a positive constant C_2 such that

$$||E||_{L_r(I)} \leq C_2 h^{\nu+1}||u||_{W_p^{\ell}(I)} \tag{4.5.3}$$

The next Lemma establishes an estimate for the first temporal derivative of the interpolation error. Initially let $D_G\sigma(\cdot)$ be the Gâteaux derivative of $\sigma(\cdot)$. Then let

$$\xi = \inf_{X \in I} \left[\frac{D_G\sigma(U_X^*)}{D_G\sigma(U_{tX}^*)} \right] \tag{4.5.4}$$

and

$$\eta = \left\| \frac{[D_G\sigma(U_X^*)],t}{D_G\sigma(U_X^*)} \right\|_{L_\infty(I)} \tag{4.5.5}$$

Then the following result holds:

<u>Lemma 4.3</u>. Let the conditions of Lemma 4.2 hold. Then there exists positive constants C_3 and C_4 independent of the mesh parameter h such that

$$||E_t||_{W_r^1(I)} \leq C_3(u_t)h^{\nu_1} ||u_t||_{W_p^m(I)} + C_4(u,u_t)h^{\nu'} ||u||_{W_p^\ell(I)} \tag{4.5.6}$$

where $\nu_1 = \min(\mu_1,\hat{\mu}_1)$, $\mu_1 = \min(k,m-1)$, $\hat{\mu}_1 = \frac{\mu_1}{r-1}$, $\nu' = \frac{\nu}{r-1}$.

<u>Proof:</u> Since σ is Gateaux differentiable on $\Omega \subset W_r^1(I)$, it can be shown (see Lemma 3.12 of [70]) that the following Lagrange formula holds:

$$((D_G\sigma(U_X^*)\,(u - \bar{z})_X, z_X))_{L_2(I)} = ((\sigma(u_X) - \sigma(\bar{z}_X), z_X))_{L_2(I)} \tag{4.5.7}$$

where

$$U_X^* = \theta u_X + (1 - \theta)\bar{z}_X \quad , \quad 0 \leq \theta \leq 1$$

Now differentiating (4.5.1) with respect to time

$$((D_G\ \sigma(U_X^*)E_{Xt}\ ,\ V_X\ ))_{L_2(I)} + (([D_G\ \sigma(U_X^*)],\ {}_tE_X,\ V_X))_{L_2(I)}$$

$$= 0 \quad \forall \quad V \in S_h^k(I) \tag{4.5.8}$$

Using (4.5.7) we get

$$(([\sigma(u_{tX}) - \sigma(\bar{z}_{tX})]\cdot \left[\frac{D_G\sigma(U_X^*)}{D_G\sigma(U_{tX}^*)}\right],\ V_X\ ))_{L_2(I)}$$

$$+ (([\sigma(u_X) - \sigma(\bar{z}_X)] \left[\frac{D_G\sigma(U_X^*),t}{D_G\sigma(U_X^*)}\right],\ V_X\ ))_{L_2(I)} = 0 \quad \forall V \in S_h^k(I) \tag{4.5.9}$$

Now let $Q \in S_h(I)$ be that element which satisfies

$$(([\sigma(u_{tX}) - \sigma(Q_{tX})] \cdot \left[\frac{D_G\sigma(U_X^*)}{D_G\sigma(U_{tX}^*)}\right],\ V_X\ ))_{L_2(I)} = 0 \quad \forall \quad V \in S_h^k(I) \tag{4.5.10}$$

We note that since the first Piola-Kirchhoff stress tensor is derived from a potential function W which we assume to be convex, ξ defined in (4.5.4) is positive. Then using an analysis similar to the one used to prove Lemma 4.1 (see Section III.3).

$$||u_t - Q_t||_{W_r^1(I)} \le C(h^{\mu_1} + \frac{C_1(u_t)}{\xi}\ h^{\hat{\mu}_1})||u_t||_{W_p^m(I)} \tag{4.5.11}$$

where $\mu_1 = \min(k,m-1)$, $\hat{\mu}_1 = \frac{\mu_1}{(r-1)}$, which can be theoretically improved to $\mu_1 - \frac{1}{2} + \frac{1}{r}$ for piecewise linear interpolation ($\mu = 1$).

Then from (4.5.9) with $V = Q_t - Z_t$

$$||Q_t - Z_t||^{r-1}_{W^1_r(I)} \leq C(u_t)||E||_{W^1_r(I)}$$

$$||Q_t - Z_t||_{W^1_r(I)} \leq C_4(u,u_t)h^{\nu'}||u||_{W^\ell_p(I)} \tag{4.5.12}$$

where $\nu' = \frac{\nu}{r-1}$. For piecewise linear case we can improve it to $\nu - \frac{1}{2} + \frac{1}{r}$, $(\mu = 1)$. In $W^1_2(I)$ norm, estimates are optimal.

Then from (4.5.11) and (4.5.12) by using the triangle inequality, we get

$$||E_t||_{W^1_r(I)} \leq C(h^{\mu_1} + \frac{C_1(u_t)}{\xi} h^{\hat{\mu}_1})||u_t||_{W^m_p(I)}$$

$$+ C_4(u,u_t)h^{\nu'}||u||_{W^\ell_p(I)}$$

for $p \geq 2$ above equation reduces to (4.5.6).

Now let

$$\nu^* = \left\| \frac{D_G\sigma(U^*_X)}{D_G\sigma(U^*_{tX})} \right\|_{L_\infty(I)}$$

Then an estimate for the first temporal derivative of the interpolation error E in the L_p norm is given in the following lemma.

Lemma 4.4. Let the hypotheses of Lemma 4.2 hold. Then there exists a positive constant C_4 such that

$$||E_t||_{L_p(I)} \le C_4\{\nu^* G(u_t,z_t)\,||E_t||_{W_p^1(I)} + \eta\, G(u,z)\,||E||_{W_p^1(I)}\}h^k$$

■

Now let

$$\overline{\xi} = \inf_{X \in I}\left[\frac{D_G\sigma(U_X^*)}{D_G\sigma(U_{tX}^*)}\right]$$

and

$$\chi = \left\|\frac{[D_G\sigma(U_X^*)]_{,t}}{D_G\sigma(U_{tX}^*)}\right\|_{L_\infty(I)}$$

$$\beta = \left\|\frac{[D_G\sigma(U_X^*)]_{,tt}}{D_G\sigma(U_X^*)}\right\|_{L_\infty(I)}$$

Then the second temporal derivative of the interpolation error in the $W_r^1(I)$ norm is given in the following Lemma:

Lemma 4.5. Let the hypotheses of Lemma 4.2 hold. Then there exists a positive constant C_5 such that

$$||E_{tt}||_{W_r^1(I)} \le C_5(u_{tt})h^{\nu_2}||u_{tt}||_{W_p^n(I)} + \hat{C}_3(u,u_t)h^{\nu'_1}||u_t||_{W_p^m(I)}$$

$$+ \hat{C}_1(u)\,h^{\nu''}||u||_{W_p^\ell(I)}$$

where $\nu_2 = \min(\mu_2,\hat{\mu}_2)$, $\mu_2 = \min(k,n-1)$, $\hat{\mu}_2 = \frac{\mu_2}{r-1}$, $\nu_1' = \frac{\nu_1}{r-1}$, $\nu'' = \frac{\nu}{(r-1)^2}$.

Proof. Direct calculation shows

$$||E_{tt}||_{W_r^1(I)} \leq C_1\left(h^{\mu_2} + h^{\hat{\mu}_2}\frac{H(u_{tt})}{\xi}\right)||u_{tt}||_{W_p^n(I)}$$

$$+ \hat{C}_3(u,u_t)[h^{\nu_1}||u_t||_{W_p^m(I)} + h^{\nu'}||u||_{W_p^{\ell}(I)}]^{1/r-1}$$

$$+ C_1(u)h^{\nu''}||u||_{W_p^{\ell}(I)}$$

In the above equation, $H(u_{tt}) = \lim_{h\to 0} G_{pp}(u_{tt},\bar{r}_{tt})$ and $\bar{r}_{tt}$ is the element in $S_h^k(I)$ such that

$$\|u_{tt} - \bar{r}_{tt}\|_{W_p^1(I)} = \inf_{r_{tt}\in S_h(I)} \|u_{tt} - r_{tt}\|_{W_p^1(I)} \quad \blacksquare$$

and if we let

$$\bar{\nu} = \left\| \frac{D_G\sigma(U_X^*)}{D_G\sigma(U_{ttX}^*)} \right\|_{L_\infty(I)}$$

Then the estimate for the second temporal derivative of the interpolative error in the L_p norm is given in the following Lemma:

Lemma 4.6. Let the hypotheses of Lemma 4.2 hold, then there exists a positive constant C_6 such that

$$\|E_{tt}\|_{L_r(I)} \leq C_6\{\bar{\nu}\, G(u_{tt},\bar{z}_{tt})\ \|E_{tt}\|_{W_p^1(I)}$$

$$+ 2\,\bar{\eta}\, G(u_t,\bar{z}_t)\ \|E_t\|_{W_p^1(I)}$$

$$+ \beta G(u,\bar{z})\ \|E\|_{W_p^1(I)}\}\ h^j \quad \blacksquare$$

The groundwork is now laid for a fairly deep theory of nonlinear Galerkin approximations.

IV.6 Accuracy and Convergence of Semidiscrete Discontinuous Finite Element/Galerkin Approximations for Wave Propagation. In this section of the report we present studies of accuracy and convergence for semidiscrete Galerkin approximations. As our primary task we wish to discuss and compare the convergence properties of the semidiscrete shock fitting scheme (4.4.12) and the semidiscrete shock smearing scheme (4.4.13). We will show by a comparison of the error characteristics for the two schemes that the conjecture of section IV.4 regarding the superiority of the shock fitting scheme in resolving dynamic response with shocks is correct.

Initially we determine the error in using the semidiscrete shock smearing scheme (4 4.13). We use a method which was developed for the linear case by Dupont [49] and which involves L_2 estimates.

Let u be the solution to (4.4.11) and U be the solution to (4.4.13). Let $e = u - U$ be the approximation error. Let Z be the element of $\overset{\circ}{M}{}^m_k(I,P)$ defined by the nonlinear energy projection

$$((\sigma(u_X) - \sigma(Z_X), V_X))_{L_2(I)} \qquad \forall \quad V \in \overset{\circ}{M}{}^m_k(I,P) \tag{4.6.1}$$

This projection is essentially a generalization of the weighted $H^1(I)$ projection introduced by Wheeler [48]. Now decompose e by setting $e = E + \mathcal{E}$ where

$$\left.\begin{aligned} E &= u - Z \\ E &= Z - U \end{aligned}\right\} \qquad (4.6.2)$$

Then the behavior of the error component E is established through the following theorem:

Theorem 4.1. Let $E = Z - U$ where Z is defined by (4.6.1) and U is the solution to (4.4.13). Let the stress operator $\sigma(\cdot)$ be Gateaux differentiable on a convex subset Ω of $W_p^1(I)$ containing u and Z. Then there exist positive constants γ and C such that

$$\| E_t \|_{L_\infty(L_2(I))} + \gamma \| E \|^{p/2}_{L_\infty(W_p^1(I))}$$

$$\leq C \{ \|E_t(0)\|_{L_2} + G(Z_X(0), U_X(0))^{1/2} \|E(0)\|_{W_p^1(I)}$$

$$+ \|E_{tt}\|_{L_2(L_2(I))} \} \qquad (4.6.3)$$

Proof: Evaluating (4.4.11) with an element $W \in \overset{\circ}{M}{}^m_k(I,P)$ and integrating by parts, we obtain

$$((\rho\ddot{u},W))_{L_2(I)} + ((\sigma(u_X),W_X))_{L_2(I)} = ((\rho f,W))_{L_2(I)} \qquad (4.6.4)$$

Similarly if (4.4.13) is evaluated for the same element $W \in \overset{\circ}{M}{}^m_k(I,P)$,

$$((\rho\ddot{U},W))_{L_2(I)} + ((\sigma(U_X),W_X))_{L_2(I)} = ((\rho f,W))_{L_2(I)} \qquad (4.6.5)$$

Subtracting (4.6.5) from (4.6.4), setting e = E + E, and identifying W with E_t

$$((\rho E_{tt}, E_t))_{L_2(I)} + ((\sigma(\mathcal{Z}_X) - \sigma(U_X), E_{Xt}))_{L_2(I)}$$

$$= ((\rho E_{tt}, E_t))_{L_2(I)} - ((\sigma(u_X) - \sigma(\mathcal{Z}_X), E_{Xt}))_{L_2(I)} \qquad (4.6.6)$$

Now since $\sigma(\cdot)$ is Gateaux differentiable the Lagrange formula [70] establishes that for some $\theta \in [0,1]$

$$((\sigma(\mathcal{Z}_X) - \sigma(U_X), E_X))_{L_2(I)} = ((D_G\sigma(U_X^*)E_X, E_X))_{L_2(I)}$$

where

$$U_X^* = \theta\mathcal{Z}_X + (1 - \theta)U_X$$

This implies that

$$((\sigma(\mathcal{Z}_X) - \sigma(U_X), E_{Xt}))_{L_2(I)} = -\frac{1}{2}(([D_G\sigma(U_X^*)]_{,t}E_X, E_X))_{L_2(I)}$$

$$+ \frac{1}{2}\frac{d}{dt}((\sigma(\mathcal{Z}_X) - \sigma(U_X), E_X))_{L_2(I)} \qquad (4.6.7)$$

Now introducing (4.6.7) into (4.6.6) and eliminating the second term on the right using the nonlinear energy projection (4.6.1)

$$\frac{1}{2}\frac{d}{dt}\left\{ \|\rho^{1/2}E_t\|^2_{L_2(I)} + ((\sigma(Z_X) - \sigma(U_X), E_X))_{L_2(I)}\right\}$$

$$\leq \frac{1}{2}(([D_G\sigma(U_X^*)]_{,t}E_X, E_X))_{L_2(I)} - ((\rho E_{tt}, E_t))_{L_2(I)} \qquad (4.6.8)$$

Using the Hölder inequality to simplify the righthand side of (4.6.8), using the inequality $ab \leq \frac{\eta}{2}a^2 + \frac{1}{2\eta}b^2$, $\eta > 0$ (henceforth to be referred to as *inequality E*), integrating from 0 to t, and using Theorems 2.6, we get

$$\|\rho^{1/2}E_t(t)\|^2_{L_2(I)} + \gamma \, \|E(t)\|^p_{W_p^1(I)}$$

$$\leq C\left\{ \|\rho^{1/2}E_t(0)\|^2_0 + \|E(0)\|^2_{W_p^1(I)}\right.$$

$$+ \|\rho^{1/2}E_{tt}\|^2_{L_2(L_2(I))} + \|\rho^{1/2}E_t\|^2_{L_2(L_2(I))}$$

$$\left. + \|[D_G\sigma(U_X^*)]_{,t}\|_{L_\infty(L_\infty(I))} \|E\|^p_{L_p(W_p^1(I))}\right\} \qquad (4.6.9)$$

The result (4.6.5) now follows by using the Lemma of Raviart [71], which gives a result very similar to the Gronwall inequality [72], and by taking the supremum over all $t \in [0,T]$ in the resulting expression.

If we next use the triangle inequality and Theorem 4.1, we obtain the final error estimate.

Theorem 4.2. Let the hypotheses of Theorem 4.1 hold. Then there exist positive constants C_1 and C_2 such that

$$\|e_t\|_{L_\infty(L_2(I))} + C_1 \|e\|^{p/2}_{L_\infty(L_p(I))}$$

$$\leq C_2 \{ \|e_t(0)\|_{L_2(I)} + \left(G(\bar{z}_X(0), U_X(0)\right)^{1/2} \|E(0)\|_{W^1_p(I)}$$

$$+ \|E\|^{p/2}_{L_\infty(L_p(I))} + \|E_t\|_{L_\infty(L_2(I))}$$

$$+ \|E_{tt}\|_{L_2(L_2(I))} \} \quad \blacksquare \qquad (4.6.10)$$

We assume that the subspace $\mathring{M}^m_k(I,P)$ has an interpolation property of order j; i.e.

$$\inf_{V \in \mathring{M}^m_k(I,P)} \|u - V\|_{W^\ell_p(I)} \leq C\, h^{j-\ell} \|u\|_{W^j_p(I)}$$

$$\text{for } \ell \leq k+1, \; 0 \leq \ell \leq m \qquad (4.6.11)$$

Then we can use the approximation theory results of Section IV.5 and Theorem 4.2 to determine the convergence characteristics of the shock smearing scheme (4.4.13). The convergence is determined by the last term in (4.6.10) ($\|E_{tt}\|_{L_2(I)}$). From Lemmas 4.1 to 4.6, we can easily deduce that

$$||E_{tt}||_{L_2(I)} \leq ||E_{tt}||_{L_p(I)}$$

$$\leq C_3[h^{\delta_2}||u_{tt}||_{W_p^{\ell}(I)} + h^{\delta_1}(||u_t||_{W_p^m(I)}$$

$$+ h^{\delta}||u||_{W_p^n(I)}]h^j$$

where $\delta = \frac{\mu}{(r-1)^3}$, $\delta_1 = \frac{\mu_1}{(r-1)^2}$, $\delta_2 = \frac{\mu_2}{(r-1)}$. Since $m = \ell+1$ and $n = \ell+2$, from the kinematical compatibility conditions, for L_2 norm,

$$||E_{tt}||_{L_2(I)} \leq Ch^j||E_{tt}||_{W_2^1(I)} \leq Ch^{j+\mu}$$

as in the linear case.

To obtain convergence to zero of the second temporal derivative of the interpolation error E we thus require that for each time point t

$$u(t), u_t(t), u_{tt}(t) \in W_p^2(I) \tag{4.6.13}$$

Thus, from (4.6.10) we find that to obtain convergence of the approximation, it is sufficient for the finite element interpolation property (4.6.11) that

$$u, u_t, u_{tt} \in L_2(W_p^2(I)) \tag{4.6.14}$$

Clearly, neither shock nor acceleration wave solutions have the regularity implied by (4.6.14). Thus, we can see that in a very weak norm (the L_2 norm in this case) the semidiscrete approximation (4.4.13) will not con-

verge to shock or acceleration wave solutions to the problem (4.3.8-4.3.9). The problem is caused by the reduced global regularity of the shock and acceleration wave solutions. Essentially anytime we attempt to define an approximation to a solution which is very irregular globally through a global projection method (of which (4.4.12) is an example), the convergence properties will deteriorate. In this case the standard interpolation property of the finite-element subspace $\overset{\circ}{M}{}^m_k(I,P)$ given by (4.6.11) is not sufficient to provide convergence to shock wave solutions. It is not a question of stability but of approximation.

We have two alternate courses of action to provide a remedy. We can increase the global regularity of the shock wave solution by introducing artificial viscosity or we can abandon the global projection idea in favor of a local projection method. In the local projection method we project onto the J_i shockless subdomains on which the exact solution has increased regularity. The embodiment of this local projection idea is the shock fitting scheme (4.4.11). We do not discuss the convergence and accuracy properties of this scheme.

In the formulation of an error estimate for the shock fitting scheme (4.4.12), we let Z be that element of $\overset{\circ}{H}{}^m(I,P,Q)$ defined by the local energy projection

$$((\sigma(u_X) - \sigma(Z_X), W_X))_{L_2(J_i)} = 0 \qquad \forall \quad W \in \overset{\circ}{H}{}^m_k(I,P,Q)$$

$$i = 1,\ldots,N$$

$$(\bar{\sigma}(u_X) - \bar{\sigma}(Z_X), \Delta W)_{Y_j} = 0 \qquad j = 1,\ldots,N-1 \qquad (4.6.15)$$

Essentially (4.6.15) represents a series of local nonlinear H_1 projections on the J_i shockless subdomains. These projections are constrained at the wavefronts Y_j by $(4.6.15)_2$ which requires that the average stress be preserved across the wave. Then let the approximation error e be decomposed into error components $e = E + \mathcal{E}$ where $E = u - Z$ and $\mathcal{E} = Z - U$.

The subspace $\overset{\circ}{H}{}^m_k(I,P,Q)$ is assumed to have the following properties:

$$\text{(i)} \quad \inf_{V \in \overset{\circ}{H}{}^m_k(I,P,Q)} \|u - V\|_{W^{\ell}_p(J_i)} \leq C\, h^{j-\ell} \|u\|_{W^j_p(J_i)}$$

$$\text{for} \quad \ell \leq k+1, \quad 0 \leq \ell \leq m, \quad i \leq 1,\ldots,N \tag{4.6.16}$$

$$\text{(ii)} \quad \|V\|_{W^j_p(J_i)} \leq k h^{-j} \|V\|_{L_p(J_i)} \qquad V \in \overset{\circ}{H}{}^m_k(I,P,Q), \quad j \leq k+1 \tag{4.6.17}$$

We do not introduce an auxiliary condition which must be applied to insure the convergence of the shock fitting scheme (4.6.12). We require that for some positive constant α an amplitude condition of the following form be satisfied:

$$\gamma \|E(t)\|^p_{W^1_p(I)} + \sum_{i=1}^{N-1} \int_0^t V_i \left\{ (\sigma^+(Z_X) - \sigma^+(U_X)) E^-_X \right.$$

$$\left. - (\sigma^-(Z_X) - \sigma^-(U_X)) E^+_X \right\} d\tau$$

$$\geq \alpha \|E(t)\|^P_{W^1_p(I)} \tag{4.6.18}$$

The second term on the lefthand side disappears in the linear case due to the Betti-Rayleigh reciprocity theorem.

The behavior of the error component E for the shock fitting scheme is given by the following theorem:

Theorem 4.3. If $E = Z - U$ where Z is defined by (4.6.16), U is defined by (4.4.12), the amplitude condition (4.6.18) is satisfied, and the intrinsic wave speed V_i, $i=1,\ldots,N-1$ is duplicated exactly by (4.4.12), then there exists constants C_1 and C_2 not depending on h such that

$$\|E_t\|_{L_\infty(L_2(I))} + C_1 \|E\|^{p/2}_{L_\infty(W^1_p(I))}$$

$$\leq C_2 \left\{ \|E_t(0)\|_{L_2(I)} + G(Z_X(0), U_X(0))^{1/2} \|E(0)\|_{W^1_p(I)} \right.$$

$$\left. + \|E_{tt}\|_{L_2(L_2(I))} + \frac{1}{h} \sup_{1 \leq i \leq N-1} V_i |||E_t|||_{L_2(Y_i(J_i, J_{i+1}))} \right\} \tag{4.6.19}$$

Proof: Evaluating (4.4.10) and (4.4.12) for the same element $W \in \overset{\circ}{H}{}^m_k(I,P,Q)$ and subtracting, we obtain

$$\sum_{i=1}^{N} ((\rho E_{tt}, W))_{L_2(J_i)} + \sum_{i=1}^{N} ((\sigma(\bar{z}_X) - \sigma(U_X), W_X))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N-1} [\, [\![\sigma(\bar{z})W]\!]_{Y_i} - [\![\sigma(U)W]\!]_{Y_i} \,]$$

$$+ \sum_{i=1}^{N-1} (\rho V(\bar{z})\, [\![\bar{z}_t]\!] - \rho V(U)\, [\![U_t]\!] + [\![\sigma(\bar{z}_X) - \sigma(U_X)]\!], \bar{W})_{Y_i}$$

$$= - \sum_{i=1}^{N} ((\rho E_{tt}, W))_{L_2(J_i)} - \sum_{i=1}^{N} ((\sigma(u_X) - \sigma(\bar{z}_X), W_X))_{L_2(J_i)}$$

$$+ \sum_{i=1}^{N-1} [\, [\![\sigma(u_X)W]\!]_{Y_i} - [\![\sigma(\bar{z}_X)W]\!]_{Y_i} \,]$$

$$- \sum_{i=1}^{N-1} (\rho V(u)\, [\![u_t]\!] - \rho V(\bar{z})\, [\![\bar{z}_t]\!] + [\![\sigma(u_X) - \sigma(\bar{z}_X)]\!], \bar{W})_{Y_i} \qquad (4.6.20)$$

Now identifying W with $E_t \in \overset{\circ}{H}{}^m_k(I,P,Q)$

$$\sum_{i=1}^{N} ((\rho E_{tt}, E_t))_{L_2(J_i)} + \sum_{i=1}^{N} ((\sigma(\bar{z}_X) - \sigma(U_X), E_{tX}))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N-1} [\, [\![\sigma(\bar{z}_X)E_t]\!]_{Y_i} - [\![\sigma(U_X)E_t]\!]_{Y_i}]$$

$$+ \sum_{i=1}^{N-1} (\rho V(\bar{z})\, [\![\bar{z}_t]\!]_{Y_i} - \rho V(U)\, [\![U_t]\!]_{Y_i} + [\![\sigma(\bar{z}_X) - \sigma(U_X)]\!], \bar{E}_t)_{Y_i}$$

$$= - \sum_{i=1}^{N} ((\rho E_{tt}, E_t))_{L_2(J_i)} - \sum_{i=1}^{N} ((\sigma(u_X) - \sigma(\bar{z}_X), E_{tX}))_{L_2(J_i)}$$

$$+ \sum_{i=1}^{N-1} [\, [\![\sigma(u_X)E_t]\!]_{Y_i} - [\![\sigma(\bar{z})E_t]\!]_{Y_i}]$$

$$- \sum_{i=1}^{N-1} (\rho V(u)\, [\![u_t]\!] - \rho V(\bar{z})\, [\![\bar{z}_t]\!]$$

$$+ [\![\sigma(u_X) - \sigma(\bar{z}_X)]\!], \bar{E}_t)_{Y_i} \tag{4.6.21}$$

Initially we use the first order kinematical condition $[\![E_t]\!]_{Y_i} = -V_i [\![E_X]\!]_{Y_i}$, which we assume to hold for the approximation U, to show that

$$[\![(\sigma(\bar{z}_X) - \sigma(U_X))E_t]\!]_{Y_i} - [\![\sigma(\bar{z}_X) - \sigma(U_X)]\!]_{Y_i} \bar{E}_{t_i}$$

$$= (\bar{\sigma}_i(Z_X) - \bar{\sigma}_i(U_X)) [\![E_t]\!]_{Y_i}$$

$$= -V_i \left(\bar{\sigma}_i(Z_X) - \bar{\sigma}_i(U_X) \right) [\![E_X]\!]_{Y_i} \tag{4.6.22}$$

Now using the Gâteaux differential in a manner similar to that presented in Theorem 4.1.

$$\frac{d}{dt} ((\sigma(Z_X) - \sigma(U_X), E_X))_{L_2(I)} = 2 \sum_{i=1}^{N} ((\sigma(Z_X) - \sigma(U_X), E_{Xt}))_{L_2(J_i)}$$

$$+ \sum_{i=1}^{N} (([D_G\sigma(U_X^*)]_{,t} E_X, E_X))_{L_2(J_i)}$$

$$+ \sum_{i=1}^{N-1} [\![(\sigma(Z_X) - \sigma(U_X)) E_X]\!]_{Y_i} V_i \tag{4.6.23}$$

where $U_X^* = \theta Z_X + (1 - \theta) U_X$.

In addition, from the definition of the L_2 norm,

$$\frac{1}{2} \frac{d}{dt} \| \rho^{1/2} E_t \|_{L_2(I)} = \sum_{i=1}^{N} ((\rho E_{tt}, E_t))_{L_2(J_i)}$$

$$+ \sum_{i=1}^{N-1} (\rho V_i [\![E_t]\!], \bar{E}_t)_{Y_i} \tag{4.6.24}$$

Introducing (4.6.22), (4.6.23), and (4.6.24) into (4.6.21) and assuming that the intrinsic velocity of the wave is duplicated exactly by the approximate model ($V_i(u) = V_i(Z) = V_i(U)$), we get

$$\frac{1}{2}\frac{d}{dt}\left\{\|\rho^{1/2}E_t\|^2_{L_2(I)} + ((\sigma(Z_X) - \sigma(U_X), E_X))_{L_2(I)}\right\}$$

$$+ \sum_{i=1}^{N-1} \frac{V_i}{2}\{(\sigma^+(Z_X) - \sigma^+(U_X))E_X^- - (\sigma^-(Z_X) - \sigma^-(U_X))E_X^+\}$$

$$= \frac{1}{2}\sum_{i=1}^{N} (([D_G\sigma(U_X^*)]_{,t}E_X, E_X))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N} ((\rho E_{tt}, E_t))_{L_2(J_i)} - \sum_{i=1}^{N} ((\sigma(u_X) - \sigma(Z_X), E_{Xt}))_{L_2(J_i)}$$

$$+ \sum_{i=1}^{N-1} 2\,(\bar{\sigma}(u_X) - \bar{\sigma}(Z_X), \Delta E_t)_{Y_i}$$

$$- \sum_{i=1}^{N-1} (\rho V_i[\![E_t]\!], \bar{E}_t)_{Y_i} \qquad (4.6.25)$$

Using the local energy projection (4.6.15) to eliminate the third and fourth terms on the righthand side and estimating the remaining terms

using the Hölder inequality and inequality E (for positive constants η and κ), we get

$$\frac{1}{2}\frac{d}{dt}\left\{\|e^{1/2}E_t\|^2_{L_2(I)} + ((\sigma(Z_X) - \sigma(U_X), E_X))_{L_2(I)}\right\}$$

$$+ \sum_{i=1}^{N-1} \frac{V_i}{2}\left\{(\sigma^+(Z_X) - \sigma^+(U_X))E_X^- - (\sigma^-(Z_X) - \sigma^-(U_X))E_X^+\right\}$$

$$\leq \frac{1}{2}\sum_{i=1}^{N} \|[D_G\sigma(U_X^*)]_{,t}\|_{L_\infty(J_i)} \|E\|^2_{W_p^1(J_i)}$$

$$+ \frac{\eta}{2}\sum_{i=1}^{N} \|\rho^{1/2}E_{tt}\|^2_{L_2(J_i)} + \frac{1}{2\eta}\sum_{i=1}^{N} \|\rho^{1/2}E_t\|^2_{L_2(J_i)}$$

$$+ \frac{\kappa}{2h^2}\sum_{i=1}^{N-1} \left|\rho^{1/2}V_i [\![E_t]\!]_{Y_i}\right|^2 + \frac{h^2}{2\kappa}\sum_{i=1}^{N-1} \left|\rho^{1/2}\bar{E}_{t_i}\right|^2 \qquad (4.6.26)$$

But $[\![E_T]\!]_{Y_i} \leq 2\, |||E_t|||_{Y_i(J_i, J_{i+1})}$ and

$$\frac{h^2}{2\kappa}\left|\rho^{1/2}\bar{E}_{t_i}\right|^2 \leq \frac{h^2}{2\kappa} \sup_{1 \leq i \leq N} \|\rho^{1/2}E_t\|^2_{L_\infty(J_i)}$$

$$\leq \frac{C^2h^2}{2\kappa} \sup_{1 \leq i \leq N} \|\rho^{1/2}E_t\|^2_{H^1(J_i)}$$

$$\leq \frac{C^2k^2}{2\kappa} \sup_{1 \leq i \leq N} \|\rho^{1/2}E_t\|^2_{L_2(J_i)} \tag{4.6.27}$$

Equation (4.6.27) follows from use of the inverse property (4.6.17) and the Sobolev imbedding theorem [3] which gives that if $y \in H^m(X)$, X is defined on an n-dimensional Euclidean space, and $m > \frac{n}{2}$, then there exists a positive constant C such that

$$\|y\|_{L_\infty(X)} \leq C \|y\|_{H^m(X)}$$

Introducing these results into (4.6.25), we get

$$\frac{1}{2}\frac{d}{dt}\left\{\|\rho^{1/2}E_t\|^2_{L_2(I)} + ((\sigma(Z_X) - \sigma(U_X), E_X))_{L_2(I)}\right\}$$

$$+ \sum_{i=1}^{N-1} \frac{V_i}{2}\left\{(\sigma^+(Z_X) - \sigma^+(U_X))E_X^- - (\sigma - (Z_X) - \sigma^-(U_X))E_X^+\right\}$$

$$\leq \frac{1}{2}\sum_{i=1}^{N} \|[D_G\sigma(U_X^*)],_t\|_{L_\infty(J_i)} \|E\|^2_{W_p^1(J_i)}$$

$$+ \frac{\eta}{2}\sum_{i=1}^{N} \|\rho^{1/2}E_{tt}\|^2_{L_2(J_i)} + \frac{1}{2\eta}\sum_{i=1}^{N} \|\rho^{1/2}E_t\|^2_{L_2(J_i)}$$

$$+ \frac{2\kappa\rho}{h^2} \sum_{i=1}^{N-1} V_i^2 |||E_t|||^2_{Y_i(J_i,J_{i+1})}$$

$$+ \frac{(N-1)C^2k^2}{2\kappa} \sup_{1 \leq i \leq N} \|\rho^{1/2} E_t\|^2_{L_2(J_i)} \tag{4.6.28}$$

Integrating from 0 to t, applying the Lemma of Raviart [71], using the response condition (4.6.18) and taking the supremum over all $t \in [0,T]$, we obtain the result (4.6.19).

The estimate for the approximation error e may be obtained by using Theorem 4.3 and the triangle inequality.

Theorem 4.4. If $e = u - U$, where u is the solution to (4.4.12), and if the amplitude condition (4.6.18) is satisfied, then there exist positive constants C_1 and C_2 not depending on h such that

$$\|e_t\|_{L_\infty(L_2(I))} + C_1 \|e\|^{p/2}_{L_\infty(L_p(I))}$$

$$\leq C_2 \Big\{ \|e_t(0)\|_0 + G(\bar{z}_x(0),U_x(0))^{1/2} \|E(0)\|_{W_p^1(I)}$$

$$+ \|E_t\|_{L_\infty(L_2(I))} + \|E\|^{p/2}_{L_\infty(L_p(I))} + \|E_{tt}\|_{L_2(L_2(I))}$$

$$+ \frac{1}{h} \sup_{1 \leq i \leq N-1} V_i |||E_t|||_{L_2(Y_i(J_i,J_{i+1}))} \Big\} \quad \blacksquare \tag{4.6.29}$$

The question of accuracy and convergence of the shock fitting scheme (4.4.12) is thus reduced to a problem in approximation theory. We must bound the interpolation error E defined by the local projection method $(4.6.15)_1$. Initially we recognize that $(4.6.15)_1$ represents a series of nonlinear elliptic boundary value problems with boundary conditions defined by $(4.6.15)_2$. This implies that we can make use of the approximation theory results of section IV.5 to define the interpolation error E on each of the shockless domains J_i.

Let $Y \in H_k^m(I,P,Q)$ be arbitrarily chosen. Then if J_k, $1 \leq k \leq N$ is a typical shockless region we have from $(4.6.15)_1$ with $W = Y - Z$ that

$$((\sigma(Y_X) - \sigma(Z_X),\ Y_X - Z_X))_{L_2(J_k)}$$

$$= ((\sigma(u_X) - \sigma(Y_X),\ Y_X - Z_X))_{L_2(J_k)} \qquad (4.6.30)$$

Then using Theorems 2.5 and 2.6 and the Holder inequality

$$\gamma \|Y_X - Z_X\|^p_{L_p(J_k)} \leq \|\sigma(u_X) - \sigma(Y_X)\|_{L_q(J_k)} \|Y_X - Z_X\|_{L_p(J_k)}$$

$$\leq G(u_X, Y_X)\| u_X - Y_X\|_{L_p(J_k)}$$

$$\|Y_X - Z_X\|_{L_p(J_k)} \qquad (4.6.31$$

Using (4.6.31), the triangle inequality, and letting $\overline{Y} \in \mathring{H}_k^m(I,P,Q)$ be that element which produces the infimum in

$$\inf_{Y \in \mathring{H}_k^m(I,P,Q)} \|u_\chi - Y_\chi\|_{L_p(J_k)}$$

then using the property (3.2.7), we get

$$\|E_\chi\|_{L_r(J_k)} \leq \inf_{Y \in \mathring{H}_k^m(I,P,Q)} \|u_\chi - Y_\chi\|_{L_p(J_k)} + \left(\frac{G(u_\chi, Y_\chi)}{\gamma}\right)^{1/r-1} \left[\inf_{Y \in \mathring{H}_k^m(I,P,Q)} \|u_\chi - Y_\chi\|_{L_p(J_k)}\right]^{1/r-1} \tag{4.6.32}$$

Using the subspace property (i) [eq. (4.6.16)] we obtain estimates for the "interpolation error" E and its temporal derivatives which are the same as developed in Lemmas 4.1 - 4.6 with the domain I replaced by J_k, $1 \leq k \leq N$. Then by the Minkowski inequality

$$\|E_\chi\|_{L_r(I)} = \left\{\sum_{k=1}^{N} \|E_\chi\|^p_{L_r(J_k)}\right\}^{1/p} \leq \sum_{k=1}^{N} \|E_\chi\|_{L_r(J_k)} \tag{4.6.33}$$

Similar relationships hold for the temporal derivatives of E.

In addition there exist positive constants C_1 and C_2 such that

$$\sup_{1 \le i \le N-1} V_i |||E_t|||_{L_2(Y_i(J_i,J_{i+1}))} \le C_1 \sum_{i=1}^{N} V_i \|E_t\|_{L_2(L_\infty(J_i))}$$

$$\le C_2 \sum_{i=1}^{N} V_i \|E_t\|_{L_2(W_p^1(J_i))} \tag{4.6.34}$$

Theorem 4.6. Let u, $u_t \in L_\infty(W_p^{k+1}(J_i))$ and $u_{tt} \in L_2(W_p^{k+1}(J_i))$ for $i = 1,\ldots,N$. Then there exist positive constants C_1 and C_2 such that

$$\|e_t\|_{L_\infty(L_2(I))} + C_1 \|e\|^{p/2}_{L_\infty(L_p(I))}$$

$$\le C_2\{||e_t(0)||_0 + G(Z_x(0),U_x(0))^{\frac{1}{2}}||E(0)||_{W_p^1(I)}$$

$$+ \sum_{i=1}^{N} h^{k+k/(r-1)^3} ||u||_{L_\infty(W_p^{k+1}(J_i))}$$

$$+ \sum_{i=1}^{N} h^{k+k/(r-1)^2} ||u_t||_{L_\infty(W_p^{k+1}(J_i))}$$

$$+ \sum_{i=1}^{N} h^{kr/(r-1)} ||u_{tt}||^{1/(r-1)}_{L_2(W_p^{k+1}(I))}$$

$$+ \sum_{i=1}^{N} h^{(k+1-r)/(r-1)} V_i ||u_t||^{1/(r-1)}_{L_2(W_p^{k+1}(J_i))} \tag{4.6.35}$$

Proof. We get this directly from Theorem 4.4, (4.6.33), (4.6.34) and Lemmas 4.1 to 4.6. ■

IV.7 Accuracy, Convergence, and Stability for Fully Discretized Models.

In this section of the paper we demonstrate the convergence of the approximation (4.4.15) to the solution to the problem (4.3.8) and (4.3.9), investigate the accuracy of the approximation (4.4.15), and discuss numerical stability criteria.

The error for the fully discretized shock fitting scheme (4.4.15) can be established through a similar procedure to the one presented in section IV.6. Initially we evaluate (4.4.10) at $t = n\Delta t$ and set $v = W \in \overset{\circ}{H}{}^m_k(I,P,Q)$. Then adding $\sum_{i=1}^{N} ((\rho\delta_t^2 u_n, W))_{L_2(J_i)}$ to each side of the resulting equation, and using the first order kinematical compatibility condition [74], we have

$$\sum_{i=1}^{N} ((\rho\delta_t^2 u^n, W))_{L_2(J_i)} + \sum_{i=1}^{N} ((\sigma(u_X^n), W_X))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N-1} [\![\sigma(u_X^n)\, W]\!]_{Y_i} + \sum_{i=1}^{N-1} (-\rho V_n^2 [\![u_X^n]\!] + [\![\sigma(u_X^n)]\!] , W)_{Y_i}$$

$$= \sum_{i=1}^{N} ((\rho f, W))_{L_2(J_i)} + \sum_{i=1}^{N} ((\varepsilon_n, W))_{L_2(J_i)} \qquad (4.7.1)$$

where u^n is the exact solution evaluated at time point $t = n\Delta t$ and

$$\varepsilon_n = \delta_t^2 u^n - \left.\frac{\partial^2 u}{dt^2}\right|_{t=n\Delta t} \tag{4.7.2}$$

We assume in this analysis that the regularity property, $\partial^4 u/\partial t^4 \in L_2(L_2(J_i))$, $i = 1,\ldots,N$ holds between shocks. Dupont [49] has shown that an estimate for ε_n is

$$\|\varepsilon_n\|^2_{L_2(J_i)} \leq C\Delta t^3 \int_{t_{n-1}}^{t_{n+1}} \left\|\frac{\partial^4 u}{dt^4}\right\|^2_{L_2(J_i)} d\tau \tag{4.7.3}$$

Setting $e_n = u^n - U^n$ and subtracting (4.4.15) from (4.7.1), we get

$$\sum_{i=1}^{N} ((\rho\delta_t^2 e_n, W))_{L_2(J_i)} + \sum_{i=1}^{N} ((\sigma(u_X^n) - \sigma(U_X^n), W_X))_{L_2(J_i)}$$

$$- \sum_{i=1}^{N-1} [\![(\sigma(u_X^n) - \sigma(U_X^n)) W]\!]_{Y_i} + \sum_{i=1}^{N-1} (-\rho V_n^2 [\![e_X^n]\!]$$

$$+ [\![\sigma(u_X^n) - \sigma(U_X^n)]\!]_Y , W)_{Y_i} = \sum_{i=1}^{N} ((\varepsilon_n, W))_{L_2(J_i)} \tag{4.7.4}$$

It is convenient to identify an element $Z^n \in \overset{\circ}{H}{}^m_k(I,P,Q)$ through the discretized local energy projection analogous to (4.6.15)

$$\left.\begin{aligned} &((\sigma(u^n_X) - \sigma(Z^n_X), W_X))_{L_2(J_i)} = 0 \qquad \forall \quad W \in \overset{\circ}{H}{}^m_k(I,P,Q) \\ &(\overline{\sigma}(u^n_X) - \overline{\sigma}(Z^n_X), \Delta W)_{Y_j} = 0 \qquad i = 1,\dots,N \\ &\qquad\qquad\qquad\qquad\qquad\qquad\quad j = 1,\dots,N-1 \end{aligned}\right\} \tag{4.7.5}$$

Then we decompose e as follows:

$$\left.\begin{aligned} e^n &= E^n + \mathcal{E}^n \\ E^n &= u^n - Z^n \\ \mathcal{E}^n &= Z^n - U^n \end{aligned}\right\} \tag{4.7.6}$$

In addition we define certain auxiliary variables by

$$\left.\begin{aligned} u_{n+1/2} &= \frac{1}{2}(u^{n+1} + u^n) \\ \delta_t\, u_{n+1/2} &= \frac{u^{n+1} - u^n}{\Delta t} \\ \delta_{t_{n+1/2}}(X) &= \frac{X|_{t=(n+1)\Delta t} - X|_{t=n\Delta t}}{\Delta t} \end{aligned}\right\} \tag{4.7.7}$$

Initially we introduce a <u>discrete</u> <u>amplitude</u> <u>condition</u> analogous to (4.6.18). We require that for positive constants γ and α

$$\gamma \|E(k\Delta t)\|^{p}_{W^1_p(I)} + \sum_{n=1}^{k+1} \sum_{i=1}^{N-1}$$

$$\Delta t \frac{V_{n_i}}{2} \{(\sigma^+(Z^n_X) - \sigma^+(U^n_X))E^{+n}_X$$

$$- (\sigma^-(Z^n_X) - \sigma^-(U^n_X))E^{-n}_X\}_{\gamma^i_n} d\tau \qquad (4.7.8)$$

$$\geq \alpha \|E(k\Delta t)\|^{p}_{W^1_p(I)}$$

$$k = 1, \ldots, r-1$$

Then the behavior of the error component E_n is given in the following Theorem:

<u>Theorem 4.7</u>. Let E_n be defined by $E_n = u^n - Z^n$ where u^n is the solution to (4.10) and Z^n is defined by (4.7.5). Let $\partial^4 u/\partial t^4$ $L_2(L_2(J_i))$ for $i = 1,\ldots,N$. In addition suppose the approximation is numerically stable, the intrinsic wave speed V_{n_i}, $i = 1,\ldots N$; $n = 1,\ldots,$ r is duplicated exactly by the approximation (4.4.15), and that the amplitude condition (4.7.8) is satisfied. Then there exist positive constants μ and ν such that

$$\|\delta_t E\|_{\tilde{L}_\infty(L_2(I))} + \mu \|E\|^{\frac{p}{2}}_{\hat{L}_\infty(W^1_p(I))}$$

$$\leq \quad \|\delta_{t\frac{1}{2}} E\|_{H^1(I)} + G(\bar{z}_X(o), U_X(o))^{1/2} \|E_o\|_{W^1_p(I)}$$

$$+ \|E_1\|_{W^1_p(I)} + \|\frac{\partial^2 E}{\partial t^2}\|_{L_2(L_2(I))}$$

$$+ \frac{1}{h} \sup_{1 \leq i \leq N-1} \hat{V}_i^2 \; |||E_X|||_{L_\infty(Y_i(J_i, J_{i+1}))}$$

$$+ \sum_{i=1}^{N} \Delta t^2 \; \|\frac{\partial^4 u}{\partial t^4}\|_{L_2(L_2(J_i))} \tag{4.7.9}$$

where

$$\|u\|_{\tilde{L}_\infty(X)} = \sup_{o \leq n \leq r-1} \|\delta_{t_{n+\frac{1}{2}}} u\|_X \; ; \; \|u\|_{\hat{L}_\infty(X)} = \sup_{o \leq n \leq r} \|u^n\|_X$$

$$\hat{V}_i = \sup_{1 \leq n \leq r} V_{n_i}$$

<u>Proof</u>: It follows from the decomposition (4.7.6) of e_n and (4.7.4) by setting $V = \delta_t E_{n+1/2} + \delta_t E_{n-1/2}$ and the assumption that the intrinsic wave speed V_n is duplicated exactly by the approximate model that

$$\sum_{i=1}^{N} ((\rho\delta_t^2 E_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}}))_{L_2(J_i^n)}$$

$$+ \sum_{i=1}^{N} ((\sigma(\bar{z}_X^n) - (U_X^n), \partial_{t_{n+\frac{1}{2}}} E_X + \delta_{t_{n-\frac{1}{2}}} E_X))_{L_2(J_i^n)}$$

$$- \sum_{i=1}^{N-1} [\![(\sigma(\bar{z}_X^n) - \sigma(U_X^n))(\delta_t E_{n+1/2} + \delta_t E_{n-1/2})]\!]_{\gamma_i^n}$$

$$+ \sum_{i=1}^{N-1} (-\rho V_n^2 [\![E_X^n]\!] + [\![\sigma(\bar{z}_X^n) - \sigma(U_X^n)]\!], \overline{\delta_t E_{n+1/2} + \delta_t E_{n-1/2}})_{\gamma_i^n}$$

$$= - \sum_{i=1}^{N} ((\rho\delta_t^2 E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_{L_2(J_i^n)}$$

$$+ \sum_{i=1}^{N} ((\sigma(u_X^n) - \sigma(\bar{z}_X^n), \delta_{t_{n+1/2}} E_X + \delta_{t_{n-1/]}} E_X))_{L_2(J_i^n)}$$

$$+ \sum_{i=1}^{N-1} [\![(\sigma(u_X^n) - \sigma(\bar{z}_X^n)(\delta_t E_{n+1/2} + \delta_t E_{n-1/2})]\!]_{\gamma_i^n}$$

$$- \sum_{i=1}^{N-1} (-\rho V_n^2 [\![E_X^n]\!] + [\![\sigma(u_X^n) - \sigma(\bar{z}_X^n)]\!], \overline{\delta_t E_{n+1/2} + \delta_t E_{n-1/2}})_{\gamma_i^n}$$

$$+ \sum_{i=1}^{N} ((\rho\varepsilon_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_{L_2(J_i^n)} \tag{4.7.10}$$

It can be shown that

$$\delta_{t_{n+1/2}} ((\sigma(\bar{Z}_X) - \sigma(U_X), E_X))_{L_2(J_i)} = ((\sigma(\bar{Z}_X^{n+1}) - \sigma(U_X^{n+1})$$

$$- \sigma(\bar{Z}_X^n) + \sigma(U_X^n), \delta_{t_{n+1/2}} E_X))_{L_2(J_i^n)}$$

$$+ 2((\sigma(\bar{Z}_X^n) - \sigma(U_X^n), \delta_{t_{n+1/2}} E_X))_{L_2(J_i^n)}$$

$$+ ((\frac{1}{\Delta t}[D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n})]E_X^{n+1}, E_X^n))_{L_2(J_i^n)}$$

$$+ \frac{1}{\Delta t}((\sigma(\bar{Z}_X^{n+1}) - \sigma(U_X^{n+1}), E_X^{n+1}))_{L_2(J_i^{n+1} - J_i^n)} \qquad (4.7.11)$$

and similarly that

$$\delta_{t_{n-1/2}} ((\sigma(\bar{Z}_X) - \sigma(U_X), E_X))_{L_2(J_i)}$$

$$= -((\sigma(\bar{Z}_X^n) - \sigma(U_X^n) - \sigma(\bar{Z}_X^{n-1}) + \sigma(U_X^{n-1}), \delta_{t_{n-1/2}} E_X))_{L_2(J_i^n)}$$

$$+ 2((\sigma(\bar{Z}_X^n) - \sigma(U_X^n), \delta_{t_{n-1/2}} E_X))_{L_2(J_i^n)}$$

$$+ \left(\left(\frac{1}{\Delta t}\,[D_G\sigma(U_X^{*n}) - D_G\sigma(U_X^{*n-1})]E_X^{n-1},\ E_X^n\right)\right)_{L_2(J_i^n)}$$

$$+ \frac{1}{\Delta t}\left(\left(\sigma(Z_X^{n-1}) - \sigma(U_X^{n-1}),\ E_X^{n-1}\right)\right)_{L_2(J_i^n - J_i^{n-1})} \qquad (4.7.12)$$

In addition another useful identity is the following:

$$\left(\left(\rho^{1/2}\,\delta_t E_{n+1/2},\ \rho^{1/2}\,\delta_t E_{n+1/2}\right)\right)_{L_2(J_i^n)}$$

$$- \left(\left(\rho^{1/2}\,\delta_t E_{n-1/2},\ \rho^{1/2}\,\delta_t E_{n-1/2}\right)\right)_{L_2(J_i^n)}$$

$$= \left(\left(\rho^{1/2}\delta_t E_{n+1/2},\ \rho^{1/2}\,\delta_t E_{n+1/2}\right)\right)_{L_2(J_i^{n+1/2})}$$

$$- \left(\left(\rho^{1/2}\,\delta_t E_{n-1/2},\ \rho^{1/2}\delta_t E_{n-1/2}\right)\right)_{L_2(J_i^{n-1/2})}$$

$$- \left(\left(\rho^{1/2}\,\delta_t E_{n+1/2},\ \rho^{1/2}\,\delta_t E_{n+1/2}\right)\right)_{L_2(J_i^{n+1/2} - J_i^n)}$$

$$- \left(\left(\rho^{1/2}\,\delta_t\,E_{n-1/2},\ \rho^{1/2}\,\delta_t\,E_{n-1/2}\right)\right)_{L_2(J_i^n - J_i^{n-1/2})}$$

(4.7.13)

Introducing (4.7.11), (4.7.12), and (4.7.13) into (4.7.10), we get

$$\frac{1}{\Delta t}\left[\sum_{i=1}^{N}((\rho^{1/2}\,\delta_t E_{n+1/2},\ \rho^{1/2}\,\delta_t E_{n+1/2}))_{L_2(J_i^{n+1/2})}\right.$$

$$\left.-\sum_{i=1}^{N}((\rho^{1/2}\,\delta_t E_{n-1/2},\ \rho^{1/2}\,\delta_t E_{n-1/2}))_{L_2(J_i^{n-1/2})}\right]$$

$$+\frac{1}{2}\sum_{i=1}^{N}\delta_{t_{n+1/2}}((\sigma(Z_X)-\sigma(U_X),E_X))_{L_2(J_i)}$$

$$+\frac{1}{2}\sum_{i=1}^{N}\delta_{t_{n-1/2}}((\sigma(Z_X)-\sigma(U_X),E_X))_{L_2(J_i)}$$

$$+\sum_{i=1}^{N-1}\frac{1}{2}V_{n_i}\,\{(\sigma^+(Z_X^n)-\sigma^+(U_X^n))E_X^{+n}$$

$$-(\sigma^-(Z_X^n)-\sigma^-(U_X^n))E_X^{-n}\}Y_i^n$$

$$+\psi^n+\beta^n+F^n+\chi^n$$

$$=-\sum_{i=1}^{N}((\rho\delta_t^2E_n,\delta_t E_{n+1/2}+\delta_t E_{n-1/2}))_{L_2(J_i^n)}$$

$$-\sum_{i=1}^{N}((\sigma(U_X^n)-\sigma(Z_X^n),\delta_{t_{n+1/2}}E_X+\delta_{t_{n-1/2}}E_X))_{L_2(J_i^n)}$$

$$+\sum_{i=1}^{N}[\![(\sigma(U_X^n)-\sigma(Z_X^n))\,(\delta_t E_{n+1/2}+\delta_t E_{n-1/2})]\!]\,Y_i^n$$

$$- \sum_{i=1}^{N-1} (-\rho V_n^2 [\![E_X^n]\!] + [\![\sigma(U_X^n) - \sigma(Z_X^n)]\!], \overline{\delta_t E_{n+1/2} + \delta_t E_{n-1/2}})_{Y_i^n}$$

$$(4.7.14)$$

$$+ \sum_{i=1}^{N} ((\rho\varepsilon_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_{L_2(J_i^n)}$$

where

$$\psi^n = - \sum_{i=1}^{N-1} (\rho V_n^2 [\![E_X^n]\!], \overline{\delta_t E_{n+1/2} + \delta_t E_{n-1/2}})_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} ((\rho^{1/2}\delta_t E_{n+1/2}, \rho^{1/2}\delta_t E_{n+1/2}))_{L_2(J_i^{n+1/2} - J_i^n)}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} ((\rho^{1/2}\delta_t E_{n-1/2}, \rho^{1/2}\delta_t E_{n-1/2}))_{L_2(J_i^n - J_i^{n-1/2})}$$

$$\beta^n = \sum_{i=1}^{N-1} \frac{V_{n_i}}{2} \{(\sigma^-(Z_X^n) - \sigma^-(U_X^n))E_X^{-n} - (\sigma^+(Z_X^n) - \sigma^+(U_X^n))E_X^{+n}\}_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} ((\sigma(Z_X^{n+1}) - \sigma(U_X^{n+1}), E_X^{n+1}))_{L_2(J_i^{n+1} - J_i^n)}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} ((\sigma(Z_X^{n-1}) - \sigma(U_X^{n-1}), E_X^{n-1}))_{L_2(J_i^n - J_i^{n-1})}$$

$$F^n = -\frac{1}{2}\sum_{i=1}^{N}((\sigma(Z_X^{n+1}) - \sigma(U_X^{n+1}) - \sigma(Z_X^n) + \sigma(U_X^n), \delta_{t_{n+1/2}}E_X))_{L_2(J_i^n)}$$

$$+\frac{1}{2}\sum_{i=1}^{N}((\sigma(Z_X^n) - \sigma(U_X^n) - \sigma(Z_X^{n-1}) + \sigma(U_X^{n-1}), \delta_{t_{n-1/2}}E_X))_{L_2(J_i^n)}$$

$$\chi^n = -\frac{1}{2}((\frac{1}{\Delta t}[D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n})]E_X^{n+1}, E_X^n))_{L_2(J_i^n)}$$

$$-\frac{1}{2}((\frac{1}{\Delta t}[D_G\sigma(U_X^{*n}) - D_G\sigma(U_X^{*n-1})]E_X^{n-1}, E_X^n))_{L_2(J_i^n)}$$

To obtain this expression we have used the fact that $[\![\delta_t E_{n+1/2} + \delta_t E_{n-1/2}]\!]_{Y_i^n} = 0$, $i = 1,\ldots,N-1$.

In the limit as $\Delta t \to 0$, ψ^n, β^n, $F^n \to 0$, however these quantities are for the discrete use (finite Δt) in general not positive. Thus they play an important role in the stability of the approximation. We proceed to estimate the size of each term. Using an approximate form of the kinematical compatibility equation $V_{n_i}[\![E_X^n]\!]_{Y_i^n} = -(\delta_t E_{n+1/2} - \delta_t E_{n-1/2})_{Y_i^n}$ we find that

$$\psi^n = \sum_{i=1}^{N-1}(\frac{\rho V_n}{2}(\delta_t E_{n+1/2} - \delta_t E_{n-1/2}), \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_{Y_i^n}$$

$$-\sum_{i=1}^{N}\frac{1}{2\Delta t}((\rho^{1/2}\delta_t E_{n+1/2}, \rho^{1/2}\delta_t E_{n+1/2}))_{L_2(J_i^{n+1/2} - J_i^n)}$$

$$-\sum_{i=1}^{N}\frac{1}{2\Delta t}((\rho^{1/2}\delta_t E_{n-1/2}, \rho^{1/2}\delta_t E_{n-1/2}))_{L_2(J_i^n - J_i^{n-1/2})} \tag{4.7.15}$$

Initially we decompose ψ^n by noting that $\psi^n = \psi_1^n + \psi_2^n$ where

$$\psi_1^n = \sum_{i=1}^{N-1} \left(\frac{\rho V}{4}^n \delta_t E_{n+1/2},\ \delta_t E_{n+1/2}\right)_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} \left(\left(\rho^{1/2}\delta_t E_{n+1/2}, \rho^{1/2}\delta_t E_{n+1/2}\right)\right)_{L_2(J_i^{n+1/2} - J_i^n)}$$

(4.7.16)

and

$$\psi_2^n = -\sum_{i=1}^{N-1} \left(\frac{\rho V n}{4} \delta_t E_{n-1/2}, \delta_t E_{n-1/2}\right)_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} \left(\left(\rho^{1/2}\delta_t E_{n-1/2}, \rho^{1/2}\delta_t E_{n-1/2}\right)\right)_{L_2(J_i^n - J_i^{n-1/2})}$$

Now we let $\delta_t E_{n+1/2}(X_i) = \delta_t E_{n+1/2}(Y_i) + X_i \frac{d}{d_{X_i}} (\delta_t E_{n+1/2}(X_i)) + \ldots +$ higher order terms. Here X_i is the normal coordinate in a coordinate system imbedded in the shock surface Y_i. Thus

$$\psi_1^n = \sum_{i=1}^{N-1} \left(\frac{\rho V n}{4} \delta_t E_{n+1/2}, \delta_t E_{n+1/2}\right)_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{\rho}{2\Delta t} \int_{J_i^{n+1/2} - J_i^n} \left[\delta_t E_{n+1/2}(Y_i) + X_i \frac{d}{d_{X_i}} (\delta_t E_{n+1/2}(X_i))\right]^2 d_X$$

$$= \sum_{i=1}^{N-1} \left(\frac{\rho V}{4}^n \delta_t E_{n+1/2}, \delta_t E_{n+1/2}\right)_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{\rho}{2\Delta t} \int_{J_i^{n+1/2} - J_i^n} \delta_t E_{n+1/2} (Y_i^2)\, d_X$$

$$- \sum_{i=1}^{N} \frac{\rho}{\Delta t} |\delta_t E_{n+1/2}|_{Y_i}^{h} \int_{J_i^{n+1/2} - J_i^n} X_i \frac{d}{X_i} (\delta_t E_{n+1/2}(X_i))dX_i$$

$$+ \ldots + \text{higher order terms} \tag{4.7.17}$$

Using the Hölder inequality, the Sobolev imbedding theorem, the approximate equation locating the singular surface $Y_i^{n+1/2}$ given by $Y_i^{n+1/2} - Y_i^n = (V_i \Delta t)/2$, and neglecting the higher order terms, we have

$$|\psi_1^n| \le \sum_{i=1}^{N} C\rho\, V_{n_i}^{3/2}\, \Delta t^{1/2}\, \|\delta_t E_{n+1/2}\|^2_{H^1(J_i^{n+1/2} - J_i^n)}$$

Using the inverse property of the subspace (4.6.17), we get

$$|\psi_1^n| \le C_1^n \frac{\Delta t^{1/2}}{h^2} \sum_{i=1}^{N} \|\delta_t E_{n+1/2}\|^2_{L_2(J_i^{n+1/2} + J_i^n)}$$

$$\le C_1^n \frac{\Delta t^{1/2}}{h^2} \|\delta_t E_{n+1/2}\|^2_{L_2(I)} \tag{4.7.18}$$

where $C_1^n = K\overline{V}_n^{3/2} N\rho C$ and $\overline{V}_n = \sup_{1 \le i \le N-1} V_{n_i}$. Through a similar calculation we can show that

$$|\psi_2^n| \le \frac{C_1^n \Delta t^{1/2}}{h^2} \|\delta_t E_{n-1/2}\|^2_{L_2(I)} \tag{4.7.19}$$

Thus (4.7.18) and (4.7.19) imply that

$$|\Delta t \sum_{n=1}^{n-1} \psi^n| \le C_1 \left[\frac{\Delta t^{3/2}}{h^2} \sum_{n=1}^{r-1} \|\delta_t E_{n+1/2}\|^2_{L_2(I)} \right.$$

$$\left. + \frac{\Delta t^{3/2}}{h^2} \sum_{n=1}^{r-1} \|\delta_t E_{n-1/2}\|^2_{L_2(I)} \right] \tag{4.7.20}$$

We can estimate β^n in a similar fashion. We rewrite (4.7.14) to obtain

$$\beta^n = \beta_1^n + \beta_2^n + \beta_3^n \tag{4.7.21}$$

where

$$\beta_1^n = \sum_{i=1}^{N-1} \frac{V_{n_i}}{2} (\sigma^-(Z_X^n) - \sigma^-(U_X^n), E_X^{-n})_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} ((\sigma(Z_X^n) - \sigma(U_X^n), E_X^n))_{L_2(J_i^{n+1} - J_i^n)}$$

$$\beta_2^n = \sum_{i=1}^{N-1} \frac{V_i^n}{2} (\sigma^+(Z_X^n) - \sigma^+(U_X^n), E_X^{+n})_{Y_i^n}$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t} ((\sigma(Z_X^n) - \sigma(U_X^n), E_X^n))_{L_2(J_i^n - J_i^{n-1})}$$

$$\beta_3^n = -\sum_{i=1}^{N} \frac{1}{2\Delta t}\left((\sigma(Z_X^{n+1}) - \sigma(U_X^{n+1}), E_X^{n+1})\right)_{L_2}(J_i^{n+1} - J_i^n)$$

$$+ \sum_{i=1}^{N} \frac{1}{2\Delta t}\left((\sigma(Z_X^{n}) - \sigma(U_X^{n}), E_X^{n})\right)_{L_2}(J_i^{n} - J_i^{n-1})$$

$$+ \sum_{i=1}^{N} \frac{1}{2\Delta t}\left((\sigma(Z_X^{n}) - \sigma(U_X^{n}), E_X^{n})\right)_{L_2}(J_i^{n+1} - J_i^{n})$$

$$- \sum_{i=1}^{N} \frac{1}{2\Delta t}\left((\sigma(Z_X^{n-1}) - \sigma(U_X^{n-1}), E_X^{n-1})\right)_{L_2}(J_i^{n} - J_i^{n-1})$$

It can be shown using a procedure similar to the one used to estimate ψ that for a positive constant C_2

$$\left|\Delta t \sum_{n=1}^{r-1} \beta^n\right| \leq C_2 \left[\frac{\Delta t^{3/2}}{h^2} \sum_{n=1}^{r-1} \|E^n\|^2_{W_p^1(I)} + \frac{\Delta t}{h^2}\left(\|E^r\|^2_{W_p^1(I)} + \|E^{r-1}\|^2_{W_p^1(I)}\right) + \|E^1\|^2_{W_p^1(I)} + \|E^0\|^2_{W_p^1(I)} \right] \quad (4.7.22)$$

Since the stress is Gateaux differentiable

$$\frac{1}{\Delta t}\left(\left(\sigma(Z_X^{n+1}) - \sigma(U_X^{n+1}) - \sigma(Z_X^n) + \sigma(U_X^n), \delta_{t_{n+1/2}} E_X\right)\right)_{L_2(J_i^n)}$$

$$= \left(\left(\frac{1}{\Delta t}\left[D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n})\right] E_X^{n+1}, \delta_{t_{n+1/2}} E_X\right)\right)_{L_2(J_i^n)}$$

$$+ \left(\left(D_G\sigma(U_X^{*n})\delta_{t_{n+1/2}} E_X - , \delta_{t_{n+1/2}} E_X\right)\right)_{L_2(J_i^n)}$$

for $U_X^{*i} = \Theta_i Z_X + (1 - \Theta_i)U_X \quad , \quad i = 1, \ldots, N.$ (4.7.23)

and

$$\frac{1}{\Delta t}\left(\left(\sigma(Z_X^n) - \sigma(U_X^n) - \sigma(Z_X^{n-1}) + \sigma(U_X^{n-1}), \delta_{t_{n-1/2}} E_X\right)\right)_{L_2(J_i^n)}$$

$$= \left(\left(\frac{1}{\Delta t}\left[D_G\sigma(U_X^{*n}) - D_G\sigma(U_X^{*n-1})\right] E_X^n, \delta_{t_{n-1/2}} E_X\right)\right)_{L_2(J_i^n)}$$

$$+ \left(\left(D_G\sigma(U_X^{*n-1})\delta_{t_{n-1/2}} E_X, \delta_{t_{n-1/2}} E_X\right)\right)_{L_2(J_i^n)} \qquad (4.7.24)$$

Introducing (4.7.23) and (4.7.24) into (4.7.14), we get

$$F_n = \frac{1}{2} \sum_{i=1}^{N} \{ - ((\left[D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n}) \right] E_X^{n+1}, \delta_{t_{n+1/2}} E_X))_{L_2(J_i^{n+1})}$$

$$+ ((\left[D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n}) \right] E_X^{n+1}, \delta_{t_{n+1/2}} E_X))_{L_2(J_i^{n+1} - J_i^n)}$$

$$- \Delta t \, ((D_G\sigma(U_X^{*n}) \delta_{t_{n+1/2}} E_X, \delta_{t_{n+1/2}} E_X))_{L_2(J_i^{n+1})}$$

$$+ \Delta t ((D_G\sigma(U_X^{*n}) \delta_{t_{n+1/2}} E_X, \delta_{t_{n+1/2}} E_X))_{L_2(J_i^{n+1} - J_i^n)}$$

$$+ ((\left[D_G\sigma(U_X^{*n}) - D_G\sigma(U_X^{*n-1}) \right] E_X^n, \delta_{t_{n-1/2}} E_X))_{L_2(J_i^n)} \} \qquad (4.7.25)$$

Now let

$$\gamma_i^n = \sup_{X \in J_i^n} |D_G\sigma(U_X^{*n})| \qquad o \le n \le r - 1$$

and

$$\xi_i^n = \sup_{X \in J_i^n} |D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n})| \quad o \le n \le r - 1$$

In addition let

$$\bar{\gamma}^n = \sup_{1 \le i \le N} |\gamma_i^n| \;\; ; \; \bar{\xi}^n = \sup_{1 \le i \le N} |\xi_i^n|$$

Multiplying F_n by Δt, summing from 1 to r-1, bounding the terms on the right hand side using the Hölder inequality and inequality E, and applying the inverse hypothesis and the Sobolev imbedding theorem, we get

$$\left|\Delta t \sum_{n=1}^{r-1} F_n\right| \leq \{ (C_3 \frac{\Delta t}{h^2} + C_4 \frac{\Delta t^2}{h^2}) \|\rho^{1/2}\delta_{t_{r-1/2}} E\|^2_{L_2(I)}$$

$$+ C_5 \Delta t \|E^r\|^2_{H^1(I)} + (C_6 \Delta t + C_7 \Delta t^2) \|\rho^{1/2}\delta_{t_{1/2}} E\|^2_{H^1(I)}$$

$$+ C_8 \Delta t \|E^1\|^2_{H^1(I)} + \frac{\Delta t^3}{h^4} \sum_{n=1}^{r-1} C_9^n \|\rho^{1/2}\delta_{t_{n+1/2}} E\|^2_{L_2(I)}$$

$$+ \frac{\Delta t^3}{h^4} \sum_{n=1}^{r-1} C_{10}^n \|E^{n+1}\|^2_{H^1(I)}\} \qquad (4.7.26)$$

where for arbiting positive constants η, β, and α

$$C_3 = \frac{k^2 \bar{\xi}^{r-1}}{4 \eta\rho} \qquad C_4 = \frac{\bar{\gamma}^{r-1}}{2\rho}$$

$$C_5 = \frac{\bar{\xi}^{r-1}\eta}{4} \qquad C_6 = \frac{\bar{\xi}^1}{4\beta\rho}$$

$$C_7 = \frac{\bar{\gamma}^1}{2\rho} \qquad C_8 = \frac{\bar{\xi}^1\beta}{4}$$

$$C_9^n = \frac{Ck^2 \bar{\xi}^n \bar{V}^n}{4\alpha\rho} + \frac{Ck^2 \bar{\gamma}^n \bar{V}^n}{2\rho}$$

$$C_{10}^n = \frac{Ck^2 \bar{\xi}^n \bar{V}_n \alpha}{4}$$

Now we assume that the Gateaux differivative of the stress has a bounded temporal derivative, i.e., for a positive constant C_{11}

$$\frac{1}{\Delta t}\left[D_G\sigma(U_X^{*n+1}) - D_G\sigma(U_X^{*n}) \right] \leq C_{11}$$

$$\frac{1}{\Delta t}\left[D_G\sigma(U_X^{*n}) - D_G\sigma(U_X^{*n-1}) \right] \leq C_{11}$$

Then it can be shown that

$$\left|\Delta t \sum_{n=1}^{r-1} \chi_n\right| \leq C_{11} \{ \Delta t \sum_{n=1}^{r-1} (\|E^{n+1}\|^2_{H^1(I)}$$

$$+ \|E^n\|^2_{H^1(I)} + \|E^{n-1}\|^2_{H^1(I)}) \} \tag{4.7.27}$$

Mutliplying (4.7.14) by Δt, summing on n from n equal 1 to r-1, and simplifying the right hand side using the discrete local energy projection (4.7.5)

$$\sum_{i=1}^{N} \|\rho^{1/2}\delta_t E_{r-1/2}\|^2_{L_2(J_i^{r-1/2})} - \sum_{i=1}^{N} \|\rho^{1/2}\delta_t E_{1/2}\|^2_{L_2(J_i^{1/2})}$$

$$+ \sum_{i=1}^{N} ((\sigma(Z_X^r) - \sigma(U_X^r), E_X^r))_{L_2(J_i^{r-1/2})}$$

$$- \sum_{i=1}^{N} ((\sigma(Z_X^{1/2}) - \sigma(U_X^{1/2}), E_X^{1/2}))_{L_2(J_i^{1/2})}$$

$$+ \sum_{n=1}^{r-1} \Delta t\psi^n + \sum_{n=1}^{r-1} \Delta t\beta^n + \sum_{n=1}^{r-1} \Delta tF^n + \Delta t \sum_{n=1}^{r-1} \chi^n$$

$$= - \sum_{n=1}^{r-1} \sum_{i=1}^{N} \Delta t ((\rho\delta_t^2 E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_{L_2(J_i^n)}$$

$$+ \sum_{n=1}^{r-1} \sum_{i-1}^{N} \Delta t(\rho V_n^2 [\![E_X^n]\!], \overline{\delta_t E_{n+1/2} + \delta_t E_{n-1/2}})_{\gamma_i^n}$$

$$+ \sum_{n=1}^{r-1} \sum_{i=1}^{N} \Delta t ((\rho\varepsilon_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_{L_2(J_i^n)} \tag{4.7.28}$$

Estimating the terms on the right hand side in (4.7.28) and using (4.7.20), (4.7.21), (4.7.26), and (4.7.25) as well as Theorems 2.5 and 2.6, we get

$$(1 - C_3 \frac{\Delta t}{h^2} - C_4 \frac{\Delta t^2}{h^2}) \|\rho^{1/2}\delta_t E_{r-1/2}\|^2_{L_2(I)}$$

$$+ (\gamma - C_2 \frac{\Delta t}{n^2} - C_5 \Delta t) \|E^r\|^p_{W_p^1(I)}$$

$$\le \{(1 + C_6\Delta t + C_7\Delta t^2)\ \|\rho^{1/2}\delta_{t_{1/2}} E\|^2_{H^1(I)}$$

$$+ G(Z_X(0),\ U_X(0))\ \|E^0\|^p_{W^1_p(I)} + C_2\ \|E^0\|^2_{W^1_p(I)}$$

$$(C_2 + C_8\Delta t)\ \|E^1\|^2_{W^1_p(I)}$$

$$+ \Delta t \sum_{n=1}^{r-1} \Big[\frac{\eta}{2}\ \|\rho^{1/2}\delta_t^2 E_n\|^2_{L_2(I)} + \sum_{i=1}^{N} \frac{\kappa}{2h^2} \left|\rho^{1/2} V^2_{n_i}\ [\![E^n_X]\!]_{Y^n_i}\right|^2$$

$$+ \sum_{i=1}^{N} \frac{h^2}{2\kappa}\ |\rho^{1/2}(\delta_t E_{n+1/2} + \delta_t E_{n-1/2})_{Y^n_i}|^2$$

$$+ \frac{\gamma}{2}\ \|\rho^{1/2}\varepsilon_n\|^2_{L_2(I)} \Big]$$

$$+ \Delta t \sum_{n=1}^{r-1} \Big[\left(\frac{1}{2\gamma} + C_9^n \frac{\Delta t^2}{h^4} + C_1 \frac{\Delta t^{1/2}}{h^2} \right) \|\rho^{1/2}\delta_{t_{n+1/2}} E\|^2_{L_2(I)}$$

$$+ \left(\frac{1}{2\gamma} + C_1 \frac{\Delta t^{1/2}}{h^2} \right) \|\rho^{1/2}\delta_{t_{n-1/2}} E\|^2_{L_2(I)}$$

$$+ \left(C_{11} + C_2 \frac{\Delta t}{h^2} + C_{10}^n \frac{\Delta t^2}{h^4}\right) \|E^{n+1}\|^2_{W_p^1(I)}$$

$$\left. + C_{11}\left(\|E^n\|^2_{W_p^1(I)} + \|E^{n-1}\|^2_{W_p^1(I)}\right)\right] \} \tag{4.7.29}$$

As <u>conditions of stability</u> we require that there exist positive constants α, β such that

$$1 - C_3 \frac{\Delta t}{h^2} - C_4 \frac{\Delta t^2}{h^2} > \alpha$$

$$\gamma - \frac{C_2 \Delta t}{h^2} - C_5 \Delta t > \alpha$$

$$1 + C_6 \Delta t + C_7 \Delta t^2 < \beta$$

$$C_2 + C_8 \Delta t < \beta$$

$$\frac{1}{2\gamma} + C_9^n \frac{\Delta t^2}{h^4} + C_1 \frac{\Delta t^{1/2}}{h^2} < \beta$$

$$\frac{1}{2\gamma} + C_1 \frac{\Delta t^{1/2}}{h^2} < \beta$$

$$C_{11} + C_2 \frac{\Delta t}{h^2} + C_{10}^n \frac{\Delta t^2}{h^4} < \beta \tag{4.7.30}$$

For stable schemes the result (4.7.9) follows by using the Sobolev imbedding theorem as in Theorem 4.4, the discrete version of the Gronwall inequality [73], and (4.7.3).

We obtain the estimate for the approximation error e^n from Theorem 4.7 and the triangle inequality.

<u>Theorem 4.8</u>. If $e^n = u^n - U^n$, where u^n is the solution to (4.4.10) and U^n is the solution to (4.4.15) and if the hypotheses of Theorem 4.7 are satisfied, then there exist positive constants ω and ϕ such that

$$\|e_t\|_{\hat{L}_\infty(L_2(I))} + \omega\|e\|^{p/2}_{\hat{L}_\infty(L_p(I))}$$

$$\leq \phi\ \{\|\delta_{t_{1/2}} E\|_{H^1(I)} + G(Z_x(o),U_x(0))^{1/2}\ \|E^\circ\|_{W^1_p(I)}$$

$$+ \|E^1\|_{W^1_p(I)} + \|E_t\|_{L_\infty(L_2(I))} + \|E\|^{p/2}_{L_\infty(L_p(I))}$$

$$+ \|E_{tt}\|_{L_2(L_2(I))} + \frac{1}{h} \sup_{1\leq i\leq N-1} V_i^2\ |||E_x|||_{L_\infty(Y_i(J_i,J_{i+1}))}$$

$$+ \sum_{i=1}^{N} \Delta t^2\ \left\|\frac{\partial^4 u}{\partial t^4}\right\|_{L_2(L_2(J_i))}\ \} \qquad (4.7.31)$$

∎

From Theorem 4.8 and Lemmas 4.1 - 4.6, we obtain the final error estimate.

Theorem 4.9. Let $u, u_t \in L_\infty(W_p^{k+1}(J_i))$, $u_{tt} \in L_2(W_p^{k+1}(J_i))$, and $u_{tttt} \in L_2(L_2(J_i))$ for $i = 1,\ldots,N$. In addition suppose that the hypotheses of Theorem 4.7 are satisfied. Then there exist positive constants ω and ζ such that

$$\|e_t\|_{\tilde{L}_\infty(L_2(I))} + \omega\|e\|^{p/2}_{\tilde{L}_\infty(L_p(I))}$$

$$\leq \zeta \; \{ \|\delta_{t_{1/2}} e\|_{H^1(I)} + G(Z_\chi(o), U_\chi(0)\; \|E^\circ\|^{p/2}_{W_p^1(I)}$$

$$+ \|E^1\|_{W_p^1(I)} + \sum_{i=1}^{N} [h^{k+k/(r-1)^3} ||u||_{L_\infty(W_p^{k+1}(J_i))}$$

$$+ h^{k+k/(r-1)^2}||u_t||^{1/(r-1)^2}_{L_\infty(W_p^{k+1}(J_i))} + h^{kr/(r-1)}||u_{tt}||^{1/(r-1)}_{L_2(W_p^{k+1}(J_i))}$$

$$+ h^{k-1+k/(r-1)^2}||u||^{1/(r-1)^2}_{L_2(W_p^{k+1}(J_i))}]\} \tag{4.7.32}$$

Now consider the practical implementation of the stability constraint (4.7.30). Initially we set

$$\gamma(\Delta t,h) = \frac{\Delta t^{1/2}}{h^2} \tag{4.7.33}$$

Then let there be a specific real number ν such that

$$\gamma \leq \nu \tag{4.7.34}$$

Then choose a Δt and select α so that

$$\alpha < 1 - C_3 \Delta t\, \nu - C_4\, \Delta t^{3/2}\, \nu^2 \tag{4.7.35}$$

We can always adjust C_5 so that this choice of α is less than $\gamma - C_2 \Delta t^{1/2}\, \nu - C_5 \Delta t$. For this choice of Δt, $C_2 + C_8 \Delta t$ can always be made less than some number β, as can $1 + \Delta t C_6 + \Delta t^2 C_7$. Thus (4.7.30), $(4.7.30)_2$, $(4.7.30)_3$, and $(4.7.30)_4$ can always be satisfied; otherwise, if $C_9^n < 0$, then $(4.7.30)_6$ is satisfied whenever $(4.7.30)_5$ is satisfied. It is thus necessary to choose ν so that $(4.7.30)_7$ is also satisfied.

In summary, we use the following procedure

1. Pick Δt.
2. Calculate the right-hand side of $(4.7.30)_3$ and $(4.7.30)_4$ Choose a specific small β that just satisfies these two inequalities.
3. Then choose ν to satisfy $(4.7.30)_5$, $(4.7.30)_6$, and $(4.7.30)_7$; i.e.,

$$\sqrt{\frac{\beta}{C_9^n} + \frac{C_1^2}{4(C_9^n)z}} - \frac{C_1}{ZC_9^n}\,,$$

$$0 < V < \min \qquad \frac{\beta - 1/2\gamma}{C_1} \tag{4.7.36}$$

$$\sqrt{\frac{\beta - C_{11}}{C_{10}^n} + \frac{(C_2 \Delta t)^2}{4(C_{10}^n)^2}} - \frac{C_2 \Delta t}{2C_{10}^n}$$

The stability constraint (4.7.30) can now be satisfied by choosing an acceptable α. A sufficient condition for stability is then to choose h so that

$$h = \frac{\Delta t^{1/4}}{\nu^{1/2}} \tag{4.7.37}$$

In the following theorem we give the stability criteria for (4.4.15) resulting from the above procedure:

Theorem 4.10. If the hypotheses of Theorem 4.7 are satisfied, then a sufficient condition to insure the numerical stability of the solution to (4.4.15) (in the $L_\infty(L_2(I))$ sense in the discrete velocities $\delta_t U_{n+1/2}$ and the $L_\infty(L_p(I))$ sense in the discrete displacements U^n) is to choose Δt and h so that (4.7.35), (4.7.36), and (4.7.37) are simultaneously satisfied.

CHAPTER V

FINITE ELEMENT IMPLEMENTATION OF THE GENERALIZED GALERKIN SHOCK FITTING TECHINQUE

V.1 Introduction. In this chapter the generalized Galerkin method, formulated and analyzed in Chapter IV. is implemented using discontinuous finite-element trial functions. In particular, we introduce finite element models for linear wave propagation, propagation of acceleration waves, and propagation of shock waves. We discuss techniques to be used for the reflection of waves, and we present various numerical examples.

V.2. Finite Element Models of Discontinuous Fields. Conventionally, finite element approximations of a function $u(X,t)$ are constructed by partitioning the domain I into a finite number E of connected subdomains over which the function is interpolated via simple polynomials. By connecting the subdomains I_e together and matching values of the local interpolants and possibly certain derivatives, at the points of intersection of the elements (i.e. the nodal points), a system of global "shape" or interpolation functions $\{\phi_\alpha(X)\}_{\alpha=1}^{E+1}$ are generated which, for each $t \in [0,T]$, form the basis of a finite dimensional space of functions. In such a construction, the approximation of the displacement field u is of the form

$$U(X,t) = \sum_{\alpha=1}^{G} A^{\alpha}(t)\Phi_{\alpha}(X) \tag{5.2.1}$$

and the velocity is

$$\dot{U}(X,t) = \sum_{\alpha=1}^{G} \dot{A}^{\alpha}(t)\Phi_{\alpha}(X) \tag{5.2.2}$$

The construction of approximate schemes for shock problems using representations such as (5.2.1) and (5.2.2) lead to an immediate paradox: $U(X,t)$ can be made continuous in X by an appropriate choice of the global basis functions $\phi_{\alpha}(X)$, but these same functions depict the behavior of the velocity $\dot{U}(X,t)$ which is known to have discontinuities at the shock front. Thus, with some deliberate modifications of the conventional finite-element Galerkin method, we will obtain schemes of the "shock-smearing" type, which depict a continuous but rapidly changing variation of velocity and displacement gradient over a shock front.

Instead of following this conventional approach, we shall construct a Galerkin scheme which contains an explicit definition of the wave front. We begin by letting P denote a partition of the particles I defined by the G nodal points

$$0 = X^0 < X^1 < \ldots < X^{G-2} < X^{G-1} = a$$

The subdomains

$$I_k = \{X\colon X^{k-1} \le X \le X^k,\ X \quad I,\ 1 \le k \le G-1\}$$

are the finite elements, $I = \sum_{k=1}^{G-1} I_k$, and we denote

$$h_k = \text{dia } I_k = X^{k-1} - X^k$$

and

$$h = \max_{0 \leq k \leq G-1} h_k$$

The real number h serves as a mesh parameter for finite element approximations.

Let $P_k(I_j)$ denote the space of polynomials of degree $\leq$ k defined on the interval I_j of I. Then a general class of <u>discontinuous finite element approximations</u> can be defined as elements of the space

$$H_h^{m,k}(I,P,Q) = \{V(X): \; V(X) \in L_2(I) \cap C^m(J_j) \cap P_k(Q_{ij}); \; 1 \leq j \leq N;$$

$$1 \leq i \leq G-1; \; Q_{ij} = I_i \cap J_j\} \tag{5.2.3}$$

Here m and k are positive integers, J_j are the shockless subdomains defined by the partition Q, I_i are the finite elements corresponding to partition P, and $P_k(Q_{ij})$ denotes the space of polynomials of degree $\leq$ k on each shockless portion of element I_i. The space $H_h^{m,k}(I,P,Q)$ is appropriate for approximating velocity fields, but we must insure that the displacement field itself remains continuous. Thus, the discontinuous finite element approximation of shock waves is a function U in $H_h^{m,k}(I,P,Q) \cap C^0(I)$, where $C^0(I)$ is the space of functions continuous on I.

When no discontinuities in the velocities or displacement gradients exist, $H_h^{m,k}(I,P,Q)$ reduces to the usual space of conforming finite element interpolants. When the discontinuities are present, the character of the discontinuity is defined within each element by a number of parameters. So as to simplify the problem of choosing and determining these parameters, we will limit ourselves to local interpolants with the following properties:

(i) The finite element basis functions $\{\Phi\} = \text{span } H_h^{m,k}(I,P,Q) \cap C^0(I)$ will have compact support in I. Indeed, for any $\Phi_\alpha(X)$ there will exist two elements I_α, $I_{\alpha+1}$ such that $\Phi_\alpha(X) = 0$ for $X \notin I_\alpha \cup I_{\alpha+1}$. Presumably, this requirement will have a beneficial effect, at least as far as the stability of the scheme is concerned.

(ii) The discontinuous basis functions vanish at the nodes of the element; on the other hand, their values at any point in an element are uniquely determined by the values of the discontinuous interpolant and possibly its derivatives on each side of the shock.

We introduce the following examples:

(i) Linear Elements. The simplest representation of a continuous function with a simple discontinuity in its first derivative is the piecewise linear local finite element representation shown in Figure 6.1a. There the wave front is located at a position Y in the element.

The piecewise linear local approximation is characterized by four parameters:

$$U(X) = \begin{cases} a + bX, & X \le Y \\ c + dX, & X > Y \end{cases} \tag{5.2.4}$$

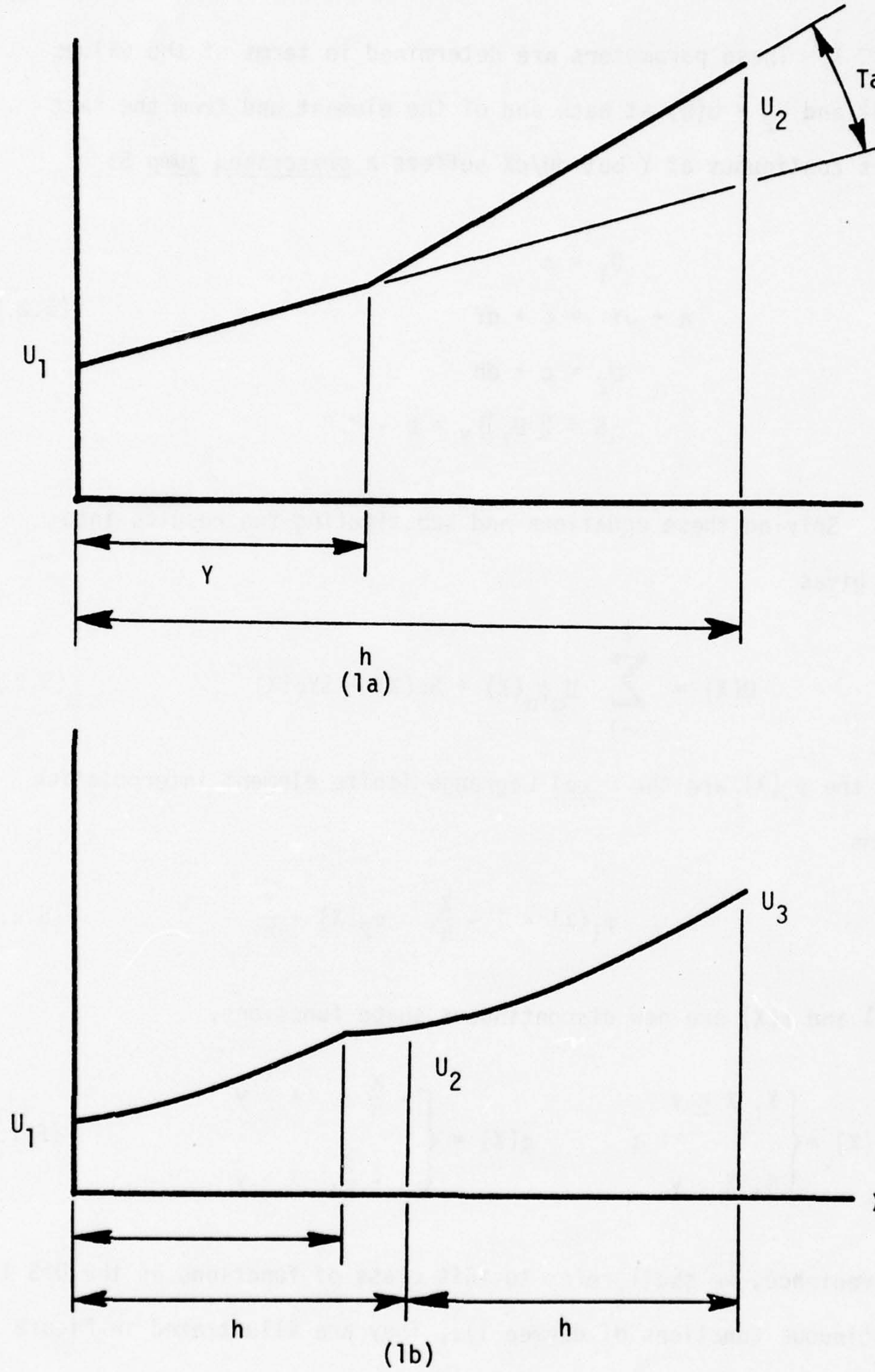


Figure 5.1 Examples of Piecewise Linear and Piecewise Quadratic Approximations.

$X \in I_\alpha \subset I$. These parameters are determined in terms of the values $U_1 = U(o)$ and $U_2 = U(h)$ at each end of the element and from the fact that U is continuous at Y but dU/dX suffers a prescribed jump S:

$$\begin{aligned} U_1 &= a \\ a + bY &= c + dY \\ U_2 &= c + dh \\ S = [\![U_X]\!]_Y &= b - d \end{aligned} \tag{5.2.5}$$

Solving these equations and substituting the results into (5.2.4) gives

$$U(X) = \sum_{\alpha=1}^{2} U_\alpha \psi_\alpha(X) + S\beta(X) + SY\phi(X) \tag{5.2.6}$$

wherein the $\psi_\alpha(X)$ are the usual Lagrange finite element interpolation functions

$$\psi_1(X) = 1 - \frac{X}{h}, \quad \psi_2(X) = \frac{X}{h} \tag{5.2.7}$$

and $\beta(X)$ and $\phi(X)$ are new discontinuous shape functions,

$$\beta(X) = \begin{cases} X, & X \leq y \\ 0, & X > y \end{cases} ; \qquad \phi(X) = \begin{cases} -\frac{X}{h}, & X \leq y \\ 1 - \frac{X}{h}, & X > y \end{cases} \tag{5.2.8}$$

For convenience, we shall refer to this class of functions as the DIS 1 (discontinuous functions of degree 1). They are illustrated in Figure 5.2.

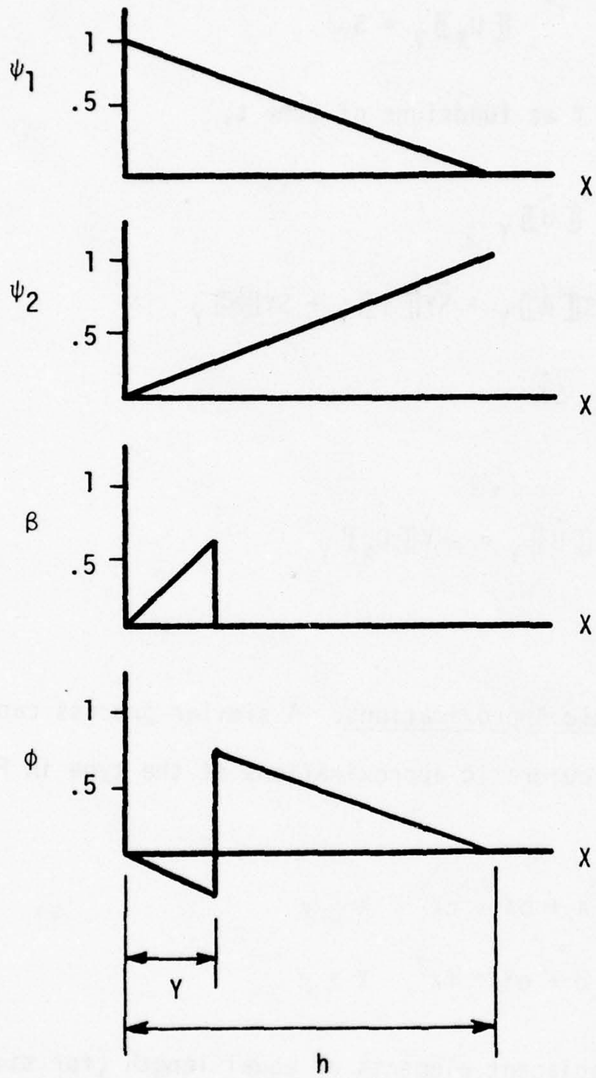


Figure 5.2 Discontinuous linear trial functions

Notice that these functions satisfy the proper kinematical compatibility conditions; i.e.

$$[\![U_X]\!]_Y = S$$

and, if we regard S and Y as functions of time t,

$$\begin{aligned}
[\![\frac{\partial U}{\partial t}]\!]_Y &= [\![\dot{U}]\!]_Y \\
&= \dot{S}[\![\beta]\!]_Y + \dot{S}Y[\![\Phi]\!]_Y + S\dot{Y}[\![\Phi]\!]_Y \\
&= - S\dot{Y}
\end{aligned}$$

Thus

$$[\![\dot{U}]\!]_Y = - \dot{Y}[\![U_X]\!]_Y \tag{5.2.9}$$

as required.

(ii) Quadratic Approximations. A similar process can be used to construct piecewise quadratic approximations of the type in Figure 5.1b. Here we use

$$U = \begin{cases} a + bX + cX^2 & X \leq y \\ d + eX + fX^2 & X > y \end{cases} \tag{5.2.10}$$

This time, we use two adjacent elements of equal length (for simplicity). Then evaluating (5.2.10) at X = o,h and 2h and denoting U(o), U(h), and U(2h) by U_1, U_2, and U_3, respectively, gives

$$\begin{aligned}
a &= U_1 \\
d + eh + fh^2 &= U_2 \\
d + 2eh + 4fh^2 &= U_3
\end{aligned} \tag{5.2.11}$$

Since U is continuous at Y

$$a + bY + cY^2 = d + eY + fY^2 \tag{5.2.12}$$

Again let S be a parameter identified with the shock strength (5.2.9) and A be a parameter to be identified with the <u>strength</u> <u>of</u> <u>the</u> <u>acceleration</u> <u>wave</u> defined by

$$A = [\![U_{XX}]\!]_Y \tag{5.2.13}$$

Thus, the definition of S and (5.2.11) - (5.2.13) are sufficient to uniquely determine the parameters in (5.2.10). We finally get

$$U(X) = \sum_{\alpha=1}^{3} U_\alpha \psi_\alpha(X) + R\beta(X) + Q\phi(X) + A\eta(X) \tag{5.2.14}$$

where $\psi_\alpha(X)$ are the usual continuous quadratic functions and the remaining terms are discontinuous quadratic interpolants:

$$\psi_1(X) = 1 - \frac{3}{2}\frac{X}{h} + \frac{X^2}{2h^2}$$

$$\psi_2(X) = 2\frac{X}{h} - \frac{X^2}{h^2}$$

$$\psi_3(X) = -\frac{X}{2h} + \frac{X^2}{2h^2}$$

$$\beta(X) = \begin{cases} X, & X \leq y \\ 0, & X > y \end{cases}$$

$$\phi(X) = \begin{cases} -\frac{3X}{2h} + \frac{X^2}{2h^2}, & X \leq Y \\ 1 - \frac{3}{2}\frac{X}{h} + \frac{X^2}{2h^2}, & X > Y \end{cases}$$

$$\eta(X) = \begin{cases} \frac{1}{2} X^2, & X \leq y \\ 0\ , & X > y \end{cases} \tag{5.2.15}$$

We shall refer to these functions as the DIS 2 family. They are illustrated in Figure 5.3.

The generalized coordinates R and Q in (5.2.14) are defined as follows

$$\left.\begin{aligned} R &= S - AY \\ Q &= Sy - \frac{1}{2} AY^2 \end{aligned}\right\} \tag{5.2.16}$$

It remains to be proved that this choice of interpolants satisfies the proper kinematical compatibility conditions at the shock front. Observe that

$$\begin{aligned} [\![U_X]\!]_Y &= (S - AY)[\![\beta_X]\!]_Y + A[\![\eta_X]\!]_Y \\ &= S - AY + AY \\ &= S \end{aligned}$$

and

$$\begin{aligned} [\![\dot{U}]\!] = (\dot{S} - \dot{A}Y - A\dot{Y})[\![\beta]\!] &+ (\dot{S}Y + S\dot{Y} \\ - \frac{1}{2}\dot{A}Y^2 - AY\dot{Y})[\![\Phi]\!]_Y &+ \dot{A}[\![\eta]\!]_Y = - S\dot{Y} \end{aligned}$$

Thus

$$[\![\dot{U}]\!]_Y = - \dot{Y}[\![U_X]\!]_Y$$

as required. This means that the <u>first</u> <u>order</u> <u>kinematical</u> <u>compatibility</u> conditions are satisfied. We shall now show that the <u>second-order</u>

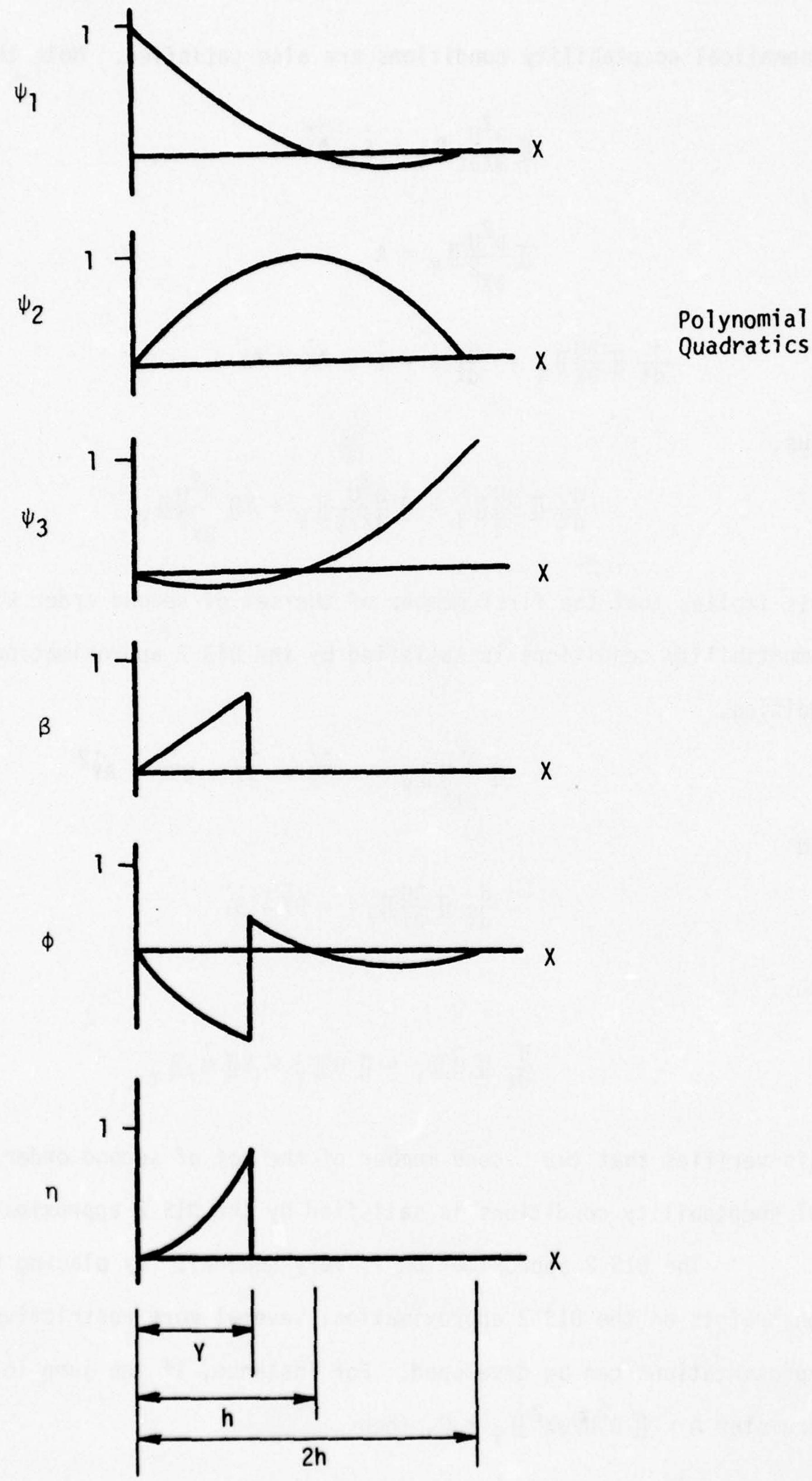


Figure 5.3 Discontinuous Quadratic Trial Functions With Discontinuity on Left Side of Internal Node.

kinematical comptability conditions are also satisfied. Note that

$$[\![\frac{\partial^2 U}{\partial X \partial t}]\!]_Y = \dot{S} - A\dot{Y}$$

$$[\![\frac{\partial^2 U}{\partial X^2}]\!]_Y = A$$

$$\frac{d}{dt} [\![\frac{\partial U}{\partial X}]\!]_Y = \frac{d}{dt} S = \dot{S} - A\dot{Y} + A\dot{Y}$$

Thus,

$$\frac{d}{dt} [\![\frac{\partial U}{\partial X}]\!]_Y = [\![\frac{\partial^2 U}{\partial X \partial t}]\!]_Y + \dot{Y} [\![\frac{\partial^2 U}{\partial X^2}]\!]_Y$$

This implies that the first member of the set of second order kinematical compatibility conditions is satisfied by the DIS 2 approximation. In addition,

$$[\![\frac{\partial^2 U}{\partial t^2}]\!]_Y = - S\ddot{Y} - \ddot{S}\dot{Y} - 2\dot{S}\dot{Y} + A\dot{Y}^2$$

and

$$\frac{d}{dt} [\![\frac{\partial U}{\partial t}]\!]_Y = - S\ddot{Y} - \dot{S}\dot{Y}$$

Thus,

$$\frac{d}{dt} [\![\dot{U}]\!]_Y = [\![\ddot{U}]\!]_Y + \dot{Y} [\![\dot{U}_X]\!]_Y$$

This verifies that the second member of the set of second order kinematical comptability conditions is satisfied by the DIS 2 approximation.

The DIS 2 approximation is very general. By placing various constraints on the DIS 2 approximation, several more restrictive representations can be developed. For instance, if the jump in the parameter $A = [\![\partial^2 U/\partial X^2]\!]_Y = 0$, then

$$U = \psi_1 U_1 + \psi_2 U_2 + \psi_3 U_3 + \beta S + \Phi SY \tag{5.2.17}$$

This will be called the DISA 2 approximation. On the other hand, if the strength of the shock wave $S = [\![\partial U/\partial X]\!]_Y = 0$, then

$$U = \psi_1 U_1 + \psi_2 U_2 + \psi_3 U_3 + \beta \overline{R} + \phi \overline{Q} + \eta A \tag{5.2.18}$$

where $\overline{R}$ and $\overline{Q}$ are generalized coordinates with

$$\overline{R} = -AY\ , \quad \overline{Q} = -\frac{1}{2} AY^2$$

This will be called the DISS 2 approximation.

An additional set of trial functions similar to the DIS 2 system is needed because the DIS 2 approximation described thus far was developed for the case in which the singular surface is on the left side of the interior node. When it is on the right side, we use the following definitions:

$$U = \psi_1 U_1 + \psi_2 U_2 + \psi_3 U_3 + \beta R + \phi Q + \eta A \tag{5.2.19}$$

$$\psi_1 = 1 - \frac{3}{2}\frac{X}{h} + \frac{X^2}{2h^2}$$

$$\psi_2 = \frac{2X}{h} - \frac{X^2}{h^2}$$

$$\psi_3 = -\frac{X}{2h} + \frac{X^2}{2h^2}$$

$$\beta = \begin{cases} 0 & X \leq X^* \\ 2h - X & X > X^* \end{cases} \tag{5.2.20}$$

$$\phi = \begin{cases} -\dfrac{X}{2h} + \dfrac{X^2}{2h^2} & X \leq X^* \\[2ex] -1 - \dfrac{X}{2h} + \dfrac{X^2}{2h^2} & X > X^* \end{cases}$$

$$\eta = \begin{cases} 0 & X \leq X^* \\ -\frac{1}{2}(2h - X)^2 & X > X^* \end{cases}$$

where

$$X^* = 2h - Y \tag{5.2.21}$$

We shall refer to this collection of functions as the DISR 2 family (R for right); they are illustrated in Figure 5.4. The first and second order kinematical compatibility equations can be shown to hold at the wavefront for the DISR 2 approximations using a method similar to the DIS 2 approximation. In addition, we shall use the notation DISRA 2 and DISRS 2 to specify the restricted approximations in which A and S are equal to zero, respectively.

<u>V.3 Models of the Particle Acceleration Field</u>. In this section we begin a discussion of the implementation of the variational principle (4.4.12) using the discontinuous trial functions introduced in V.2. We emphasize here the linear trial functions (the DIS 1 set). In particular we consider here models of the acceleration field.

For convenience in notation we define a new trial function $\chi = \beta + Y\phi$. Then the displacement approximation corresponding to (5.2.6) is

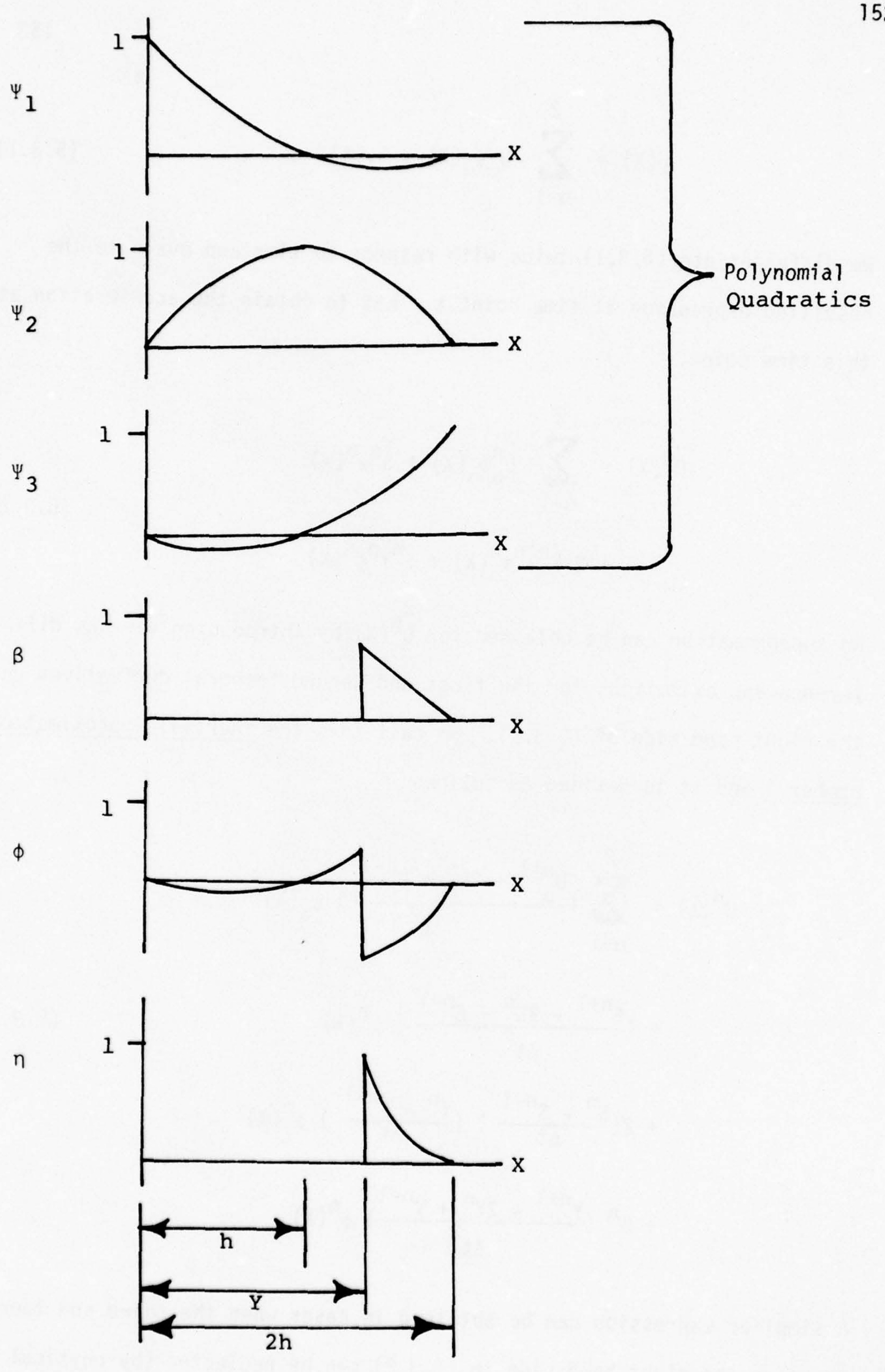


FIGURE 5.4 Discontinuous Quadratic Trial Functions With Discontinuity on Right Side of Internal Node

$$U(X) = \sum_{\alpha=1}^{2} U_\alpha \psi_\alpha(X) + S\chi(X) \tag{5.3.1}$$

We differentiate (5.3.1) twice with respect to time and evaluate the resulting expression at time point $t = n\Delta t$ to obtain the acceleration at this time point.

$$\ddot{U}^n(X) = \sum_{\alpha=1}^{2} \ddot{U}^n_\alpha \psi_\alpha(X) + \ddot{S}^n \chi^n(X) + 2\,\dot{S}^n \dot{Y}^n \phi^n(X) + S^n \ddot{Y}^n \phi^n(X) \tag{5.3.2}$$

An approximation can be obtained for $\ddot{U}^n(X)$ by introducing various difference approximations for the first and second temporal derivatives on the right hand side of (5.3.2). We call this the <u>inertial approximation number 1</u> and it is defined as follows:

$$\begin{aligned} \ddot{U}^n(X) = & \sum_{\alpha=1}^{2} \left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) \psi_\alpha(X) \\ & + \left(\frac{S^{n+1} - 2S^n + S^{n-1}}{\Delta t^2}\right) \chi^n(X) \\ & + 2\left(\frac{S^n - S^{n-1}}{\Delta t}\right)\left(\frac{Y^n - Y^{n-1}}{\Delta t}\right) \phi^n(X) \\ & + S^n \left(\frac{Y^{n+1} - 2Y^n + Y^{n-1}}{\Delta t^2}\right) \phi^n(X) \end{aligned} \tag{5.3.3}$$

A simplier expression can be obtained in cases when the third and fourth terms on the right hand side in (5.3.2) can be neglected (by physical

arguments) compared to the first two terms on the right hand side. We call this inertial approximation number 2. It is defined as follows:

$$\ddot{U}^n(X) = \sum_{\alpha=1}^{2} \left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) \psi_\alpha(X)$$

$$+ \left(\frac{S^{n+1} - 2S^n + S^{n-1}}{\Delta t^2}\right) \chi^n(X) \tag{5.3.4}$$

A third alternative acceleration term can be obtained by differencing $\ddot{U}(X)$ directly using central differences. This we call inertial approximation number 3. It is defined by the following expression:

$$\ddot{U}^n(X) = \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2}$$

$$= \sum_{\alpha=1}^{2} \left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) \psi_\alpha(X) \tag{5.3.5}$$

$$+ \frac{S^{n+1}\chi^{n+1}(X) - 2S^n\chi^n(X) + S^{n-1}\chi^{n-1}(X)}{\Delta t^2}$$

Several approximation schemes can be developed from these alternative inertial formulations. We will discuss these alternate approximation schemes in the following sections.

V.4 An Unsuccessful Model. Initially we consider a model which uses inertial approximation number 1. We essentially use the formulation presented in Chapter V (in particular equation 4.4.12) except that we pose the equations on the element level. We consider a one dimensional finite element I_e which is partitioned into two disjoint shockless

domains J_{e_1} and J_{e_2} as shown in Figure 5.5. We introduce (5.3.2) into (5.4.12) and select W from the set $\{\psi_1, \psi_2, \beta, \phi\}$. The following system of equations is obtained:

$$\sum_{\alpha=1}^{2} M_{\alpha\beta}\left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) + B_\beta^n \left(\frac{S^{n+1} - 2S^n + S^{n-1}}{\Delta t^2}\right)$$

$$+ C_\beta^n S^n\left(\frac{Y^{n+1} - 2Y^n + Y^{n-1}}{\Delta t^2}\right) + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \psi_{\beta,X} dX$$

$$= -2C_\beta^n\left(\frac{S^n - S^{n-1}}{\Delta t}\right)\left(\frac{Y^n - Y^{n-1}}{\Delta t}\right) \tag{5.4.1}$$

$$+ \rho V_n^2 \, S^n \bar{\psi}_\beta(Y^n) + f_\beta + S_\beta^n$$

$$\sum_{\alpha=1}^{2} D_\alpha^n\left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) + E^n \left(\frac{S^{n-1} - 2S^n + S^{n-1}}{\Delta t^2}\right)$$

$$+ F^n S^n\left(\frac{Y^{n+1} - 2Y^n + Y^{n-1}}{\Delta t^2}\right) + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \beta^n_{,X} \, dX \tag{5.4.2}$$

$$= -2F^n\left(\frac{S^n - S^{n-1}}{\Delta t}\right)\left(\frac{Y^n - Y^{n-1}}{\Delta t}\right)$$

$$+ \rho V_n^2 S^n \bar{\beta}^n + \bar{\sigma}^n [\![\beta^n]\!]_{Y^n} + R^n$$

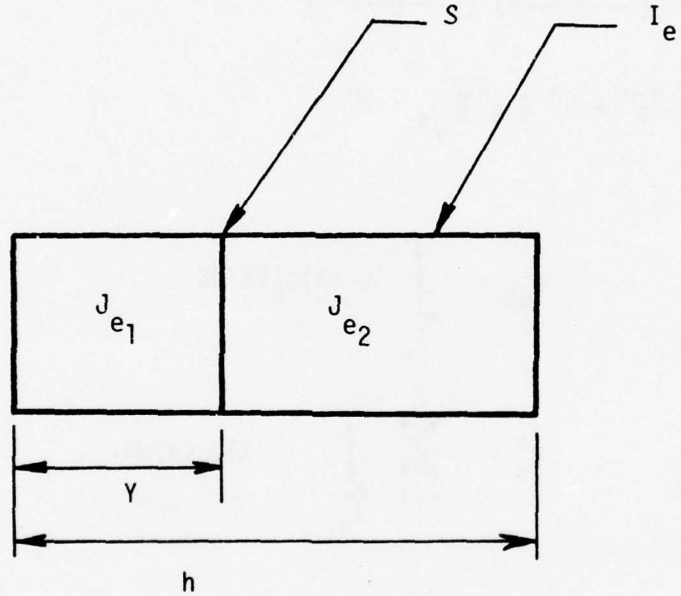


Figure 5.5 One-Dimensional Finite Element With a Singular Surface S.

$$\sum_{\alpha=1}^{2} H_\alpha^n \left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) + Q^n\left(\frac{S^{n+1} - 2S^n + S^{n-1}}{\Delta t^2}\right)$$

$$+ T^n S^n\left(\frac{Y^{n+1} - 2Y^n + Y^{n-1}}{\Delta t^2}\right) + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \phi^n_{,X}\, dX$$

$$= - 2T^n\left(\frac{S^n - S^{n-1}}{\Delta t}\right)\left(\frac{Y^n - Y^{n-1}}{\Delta t}\right) \tag{5.4.3}$$

$$+ \rho V_n^2 S^n \overline{\phi}^n + \overline{\sigma}^n [\![\phi^n]\!]_{Y^n} + Z^n$$

where

$$M_{\alpha\beta} = \int_{I_e} \psi_\alpha(X)\psi_\beta(X)dX$$

$$B_\beta^n = \sum_{i=1}^{2} \int_{J_{e_i}^n} \chi^n(X)\psi_\beta(X)dX$$

$$C_\beta^n = \sum_{i=1}^{2} \int_{J_{e_i}^n} \phi^n(X)\psi_\beta(X)dX$$

$$D_\alpha^n = \sum_{i=1}^{2} \int_{J_{e_i}^n} \psi_\alpha(X)\beta^n(X)dX$$

$$E^n = \sum_{i=1}^{2} \int_{J_{e_i}^n} \chi^n(X)\beta^n(X)dX$$

$$F^n = \sum_{i=1}^{2} \int_{J^n_{e_i}} \phi^n(X)\beta^n(X)dX$$

$$H^n_\alpha = \sum_{i=1}^{2} \int_{J^n_{e_i}} \psi_\alpha(X)\phi^n(X)dX$$

$$Q^n = \sum_{i=1}^{2} \int_{J^n_{e_i}} \chi^n(X)\phi^n(X)dX$$

$$T^n = \sum_{i=1}^{2} \int_{J^n_{e_i}} \phi^n(X)\phi^n(X)dX$$

$$f_\beta = \int_{I_e} f\psi_\beta dX$$

$$S^n_\beta = \sigma^n\psi_\beta\Big|_o^{L_e}$$

$$R^n = \sum_{i=1}^{2} \int_{J^n_{e_i}} f\ \beta^n dX$$

$$Z^n = \sum_{i=1}^{2} \int_{J^n_{e_i}} f\phi^n dX$$

We normally set

$$V_n = \frac{Y^n - Y^{n-1}}{\Delta t} \tag{5.4.4}$$

This scheme is started with a forward difference approximation at time point t = 0. We require the following conditions to define the initial data and starting procedure:

$$\int_{I_e} [\sum_{\alpha=1}^{2} U_\alpha^o \psi_\alpha(X) + S^o \chi^o(X)]\psi_\beta(X)dX \tag{5.4.5}$$

$$= \int_{I_e} u(X,0)\psi_\beta(X)dX$$

$$\int_{I_e} [\sum_{\alpha=1}^{2} U_\alpha^o \psi_\alpha(X) + S^o \chi^o(X)]\chi^o(X)dX$$

$$= \int_{I_e} u(X,0)\chi^o(X)dX \tag{5.4.6}$$

$$\int_{I_e} [\sum_{\alpha=1}^{2} (\frac{U_\alpha^1 - U_\alpha^o}{\Delta t})\psi_\alpha(X) + (\frac{S^1 - S^o}{\Delta t})\chi^o(X)$$

$$+ S^o V_o \phi^o(X)]\psi_\beta(X)dX \tag{5.4.7}$$

$$= \int_{I_e} \dot{u}(X,0)\psi_\beta(X)dX$$

$$\int_{I_e} [\sum_{\alpha=1}^{2} (\frac{U^1_\alpha - U^0_\alpha}{\Delta t})\psi_\alpha(X) + (\frac{S^1 - S^0}{\Delta t})\ \chi^0(X)$$

$$+ S^0 V_0 \phi^0(X)]\chi^0(X)dX \tag{5.4.8}$$

$$= \int_{I_e} \dot{U}(X,0)\chi^0(X)dX$$

$$V_0^2 = \frac{[\![\sigma^0]\!]\ Y^0}{\rho S^0} \tag{5.4.9}$$

This scheme (5.4.1 - 5.4.4) will be characterized as an <u>explicit shock position method</u>. The position of the shock is a dependent variable. Essentially we integrate the second time derivative of the shock position $\ddot{Y}$ twice to obtain the change in position of the shock. Presumably this formulation should allow for more accurate determination of the shock position and should more consistently bring the shock position into the formulation since it appears as a dependent variable. Unfortunately the <u>explicit shock position method</u> appears to be unstable in calculations. We do not know why the scheme is unstable; however, we suspect that a more judcious choice of the trial functions and the approximation for the temporal operator will produce a stable scheme. At any rate we have discussed it here because this approximation seems to be the most natural, and we hope that future computations can improve its performance.

V.5 Equations of Motion using Finite Elements with Discontinuous Fields. We can obtain another shock fitting method by using inertial approximation number 2 (equation 5.3.4). Introducing (5.3.4) into (4.4.12) and selecting W from $\{\psi_1, \psi_2, \beta\}$, we get

$$\sum_{\alpha=1}^{2} M_{\alpha\beta}\left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) + B_\beta^n\left(\frac{S^{n+1} - 2S^n + S^{n-1}}{\Delta t^2}\right) + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \psi_{\beta,X}\, dX = \rho V_n^2 S^n \overline{\psi}_\beta (Y^n) + f_\beta + S_\beta^n \tag{5.5.1}$$

$$\sum_{\alpha=1}^{2} D_\alpha^n\left(\frac{U_\alpha^{n+1} - 2U_\alpha^n + U_\alpha^{n-1}}{\Delta t^2}\right) + E^n\left(\frac{S^{n+1} - 2S^n + S^{n-1}}{\Delta t^2}\right) + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \beta^n_{,X}\, dX = \rho V_n^2 S^n \overline{\beta}^n + \overline{\sigma}^n [\![\beta^n]\!]_{Y^n} + R^n \tag{5.5.2}$$

An additional equation is required to define Y^n. We set

$$\frac{Y^{n+1} - Y^n}{\Delta t} = V_n \tag{5.5.3}$$

where

$$V_n^2 = \frac{[\![\sigma^n]\!]_{Y^n}}{\rho S^n} \tag{5.5.4}$$

(5.5.4) is an expression for the shock speed evaluated at time point $t = n\Delta t$. We call (5.5.1 - 5.5.4) an implicit shock position method. The initial conditions and starting procedure for this method are the same as (5.4.5 - 5.4.9).

In the implicit shock position method the velocity of the shock is defined by a relationship (5.5.4) valid only at the wave front and only involves the data at time point $t = n\Delta t$. In the explicit shock position method the velocity of the shock is defined by integrating the acceleration of the shock. The acceleration of the shock is defined in turn by global balance laws and is determined from the solution at time point $t = n\Delta t$ and time point $t = (n-1)\Delta t$. Thus there is a significant difference between the two approximations. As mentioned in V.4 the explicit shock position method seems to be unstable for the specific choice of trial functions and integration algorithm used there. However the implicit shock fitting method turns out to be stable for reasonable choices of the discretization parameters. Thus it has been our main tool in numerical calculations.

An assumption in the implicit shock fitting method is that the last two terms on the right hand side in (5.3.2) are small compared to the first two. In cases in which this assumption cannot be justified we can still construct an implicit shock fitting method. This new method is based on inertial approximation number 3. Introducing (5.3.5) into (4.4.12) and selecting W from $\{\psi_1, \psi_2, \beta\}$. we get

$$\sum_{\alpha=1}^{2} M_{\alpha\beta}\left(\frac{U_\alpha^{n+1} - 2U_\alpha^{n} + U_\alpha^{n-1}}{\Delta t^2}\right) + \frac{1}{\Delta t^2} B_\beta^{n+1} S^{n+1}$$

$$- \frac{2}{\Delta t^2} B_\beta^{n} S^{n} + \frac{1}{\Delta t^2} B_\beta^{n-1} S^{n-1} + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \psi_{\beta,X} dX \tag{5.5.5}$$

$$= \rho V_n^2 S^n \overline{\psi}_\beta(Y^n) + f_\beta + S_\beta^n$$

$$\sum_{\alpha=1}^{2} D_\alpha^n\left(\frac{U_\alpha^{n+1} - 2U_\alpha^{n} + U_\alpha^{n-1}}{\Delta t^2}\right) + \frac{1}{\Delta t^2} K^n S^{n+1}$$

$$- \frac{2}{\Delta t^2} E^n S^n + \frac{1}{\Delta t^2} L^n S^{n-1} + \sum_{i=1}^{2} \int_{J_{e_i}^n} \sigma^n \beta^n_{,X} dX \tag{5.5.6}$$

$$= \rho V_n^2 S^n \overline{\beta}^n + \overline{\sigma}^n [\![\beta^n]\!]_{Y^n} + R^n$$

$$\frac{Y^{n+1} - Y^n}{\Delta t} = V_n \tag{5.5.7}$$

$$V_n^2 = \frac{[\![\sigma^n]\!]_{Y^n}}{\rho S^n} \tag{5.5.8}$$

where

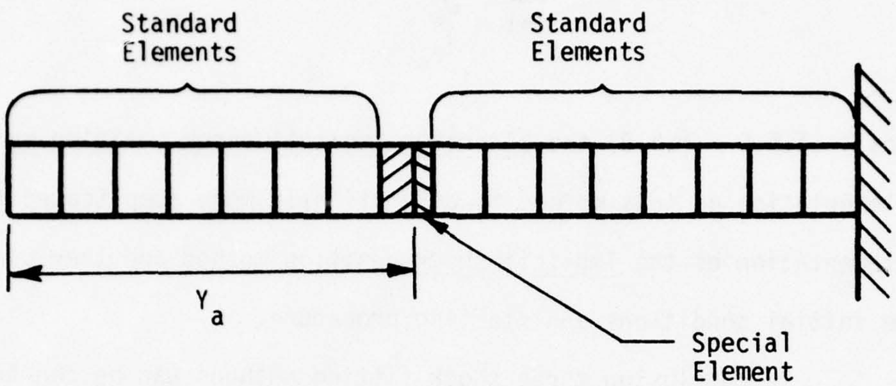
$$K^n = \sum_{i=1}^{2} \int_{J_{e_i}^n} \chi^{n+1}(X) \beta^n(X) dX$$

$$L^n = \sum_{i=1}^{2} \int_{J^n_{e_i}} \chi^{n-1}(X)\beta^n(X)dX$$

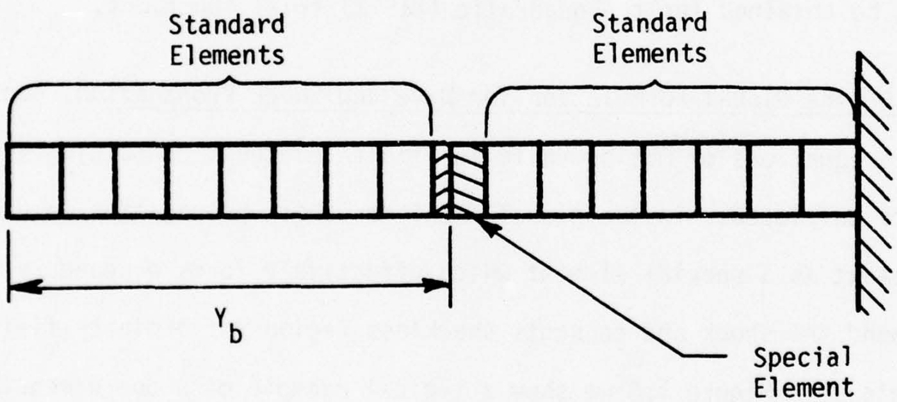
We call (5.5.5 - 5.5.8) the alternate implicit shock position method. Implementation of this method is only slightly more complicated than the implementation of the implicit shock position method and they use the same initial conditions and starting procedure.

In conclusion three shock fitting methods can be constructed from the variational principle (4.4.12). Each of these approximations has advantages in specific cases. Specific algorithms have been presented for the linear (DIS 1) trial functions. Analogous approximations can be obtained for the quadratic (DIS 2) trial functions.

V.6 The Global Formulation for Wave and Shock Propagation. In section V.5 equations of motion valid for finite elements containing shock waves were developed. In the global model for shock propagation we use this element as a special element which effectively forms a boundary layer around the shock and connects shockless regions of ordinary finite elements. In Figure 5.6 we show a typical example of a one dimensional bar containing a single shock wave. The bar is divided into 16 finite elements whose nodes correspond to fixed material coordinates; i.e., we use a Lagrangian coordinate system for the elements. The shock wave forms a moving surface in the element designated as the special element. All other elements in the model are standard. When the shock surface crosses inner-element boundaries the special element also moves and the previous special element reverts to a standard finite element. In Figure 5.6 the



(a) t = a



(b) t = b

Figure 5.6 Finite Element Model with Special Element as a Boundary Layer at the Singular Surface.

finite element model using the DIS 1 approximation is shown at a time point which we designate as $t = a$ and at some later time point which we designate as $t = b$. The global basis functions associated with the special element for these two time points are shown in Figure 5.7. This demonstrates the change in the finite element model as the shock crosses inner-element boundaries. In Figure 5.8 we show the change in the global basis functions for the special element as the shock crosses the center of the special element.

Essentially the $\beta(X)$, $\phi(X)$, and $\chi(X)$ trial functions change at each time point. We have not used any remapping procedure, and we do not feel that it is necessary since this variation in the trial functions is effectively accounted for in the equation of motion.

The approximation has certain peculiarities which are given in the following comments:

i) The mass matrix associated with the special element is not diagonal and is not symmetric. The inverse mass matrix for the special element is fully populated. Thus the first node in front of a shock wave usually is accelerated even if in the exact solution the region in front of the wave is unstrained. As the mesh is refined, the loading of the region in front of the wave disappears. Concerning the inversion of the mass matrix, if a diagonal mass matrix is used for all standard elements then the mass matrix for the special element can be inverted separately and superimposed on the inverse matrix resulting from the diagonal portion. Thus, there is no computational disadvantage because of the special element.

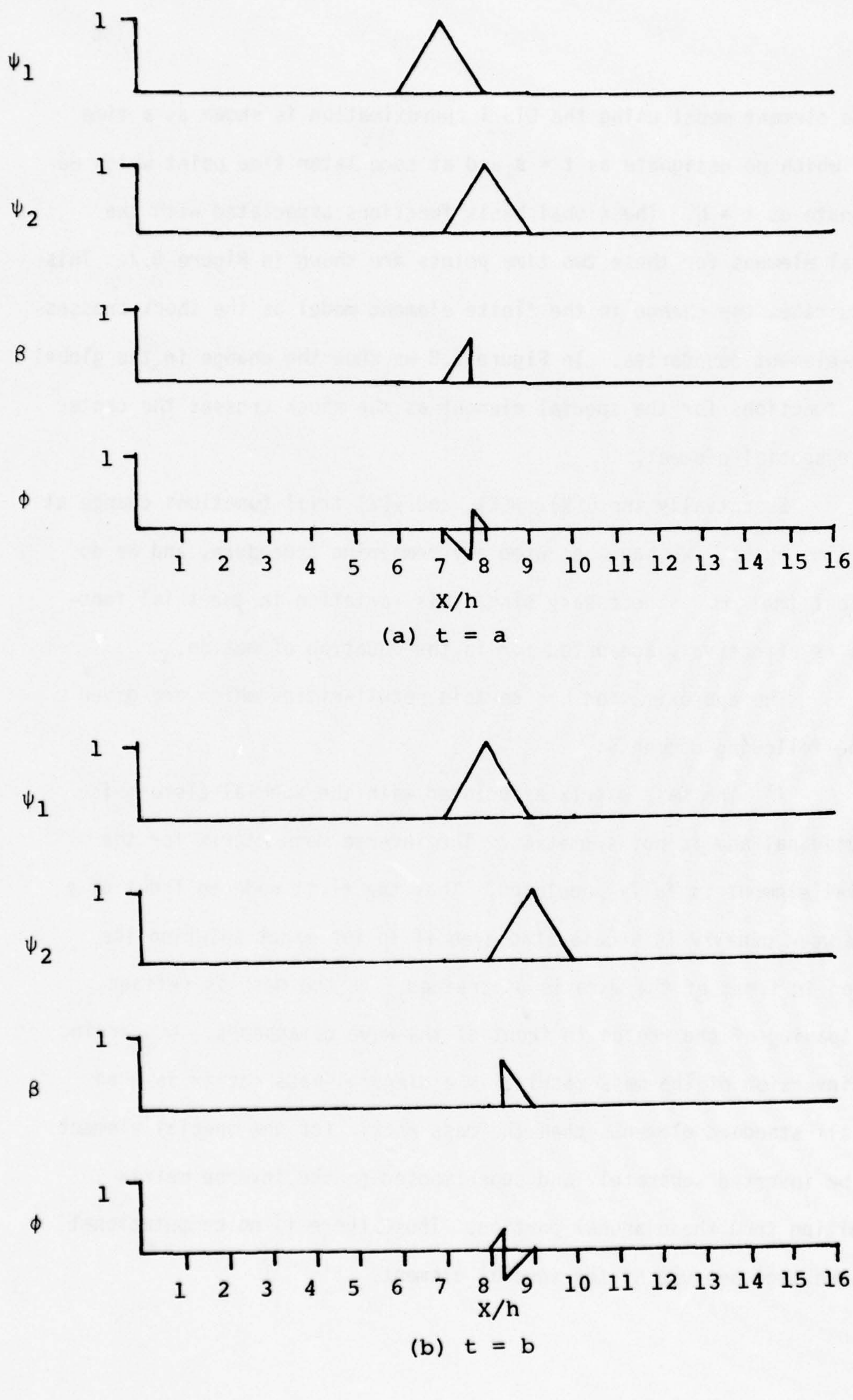
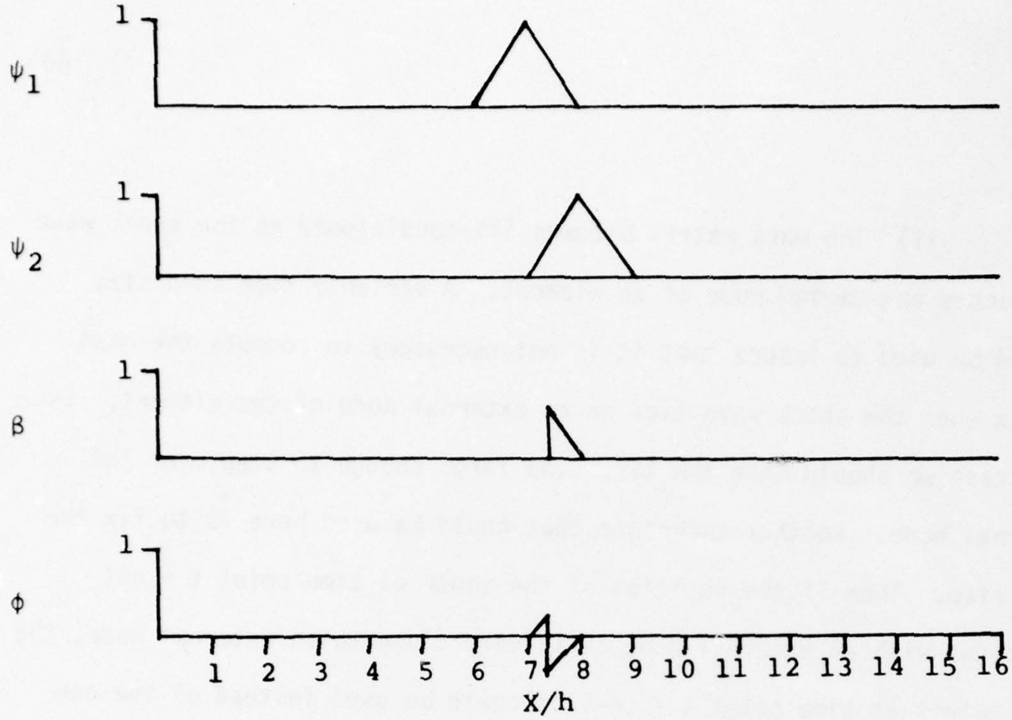
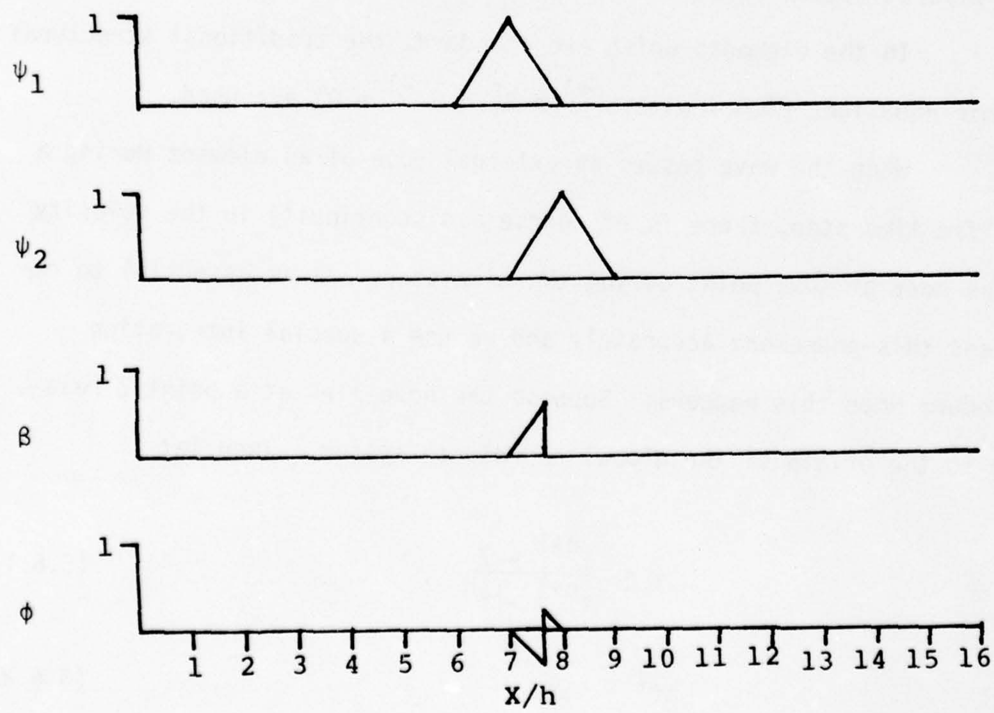


FIGURE 5.7 Change in Global Basis Functions Corresponding to the Special Element on Crossing Inner-Element Boundaries



(a) Discontinuity on Left Side of Element



(b) *Discontinuity on Right Side of Element*

FIGURE 5.8 Change in Global Basis Functions Corresponding to the Special Element in Crossing Mid-Element

ii) The mass matrix becomes ill-conditioned as the shock wave approaches an external node of an element. A variable time step size should be used to insure that it is not necessary to compute the mass matrix when the shock wave lies on an external node of the element. In this case we should make the time step large enough to step over the external node. Another technique that could be used here is to fix the time step. Then if the position of the shock at time point $t = n\Delta t$ coincides with an external node or is very close to an external node, the mass matrix at time point $t = (n-1)\Delta t$ could be used instead of the one associated with time point $t = n\Delta t$. This effect also disappears as the mesh discretization parameters are contracted.

In the elements which are standard, the traditional structural dynamic equations (5.4.1 with $S^{n+1} = S^n = S^{n-1} = 0$) are used.

When the wave passes an external node of an element during a specific time step, there is of course a discontinuity in the velocity of the node at some point during the time step. It is essential to represent this phenomena accurately and we use a special integration procedure when this happens. Suppose the node lies at a point Z relative to the origin of the global coordinate system. Then let

$$\gamma = \frac{Y^{n+1} - Z}{Y^{n+1} - Y^n} \qquad (5.6.1)$$

$$\Delta t' = \gamma \Delta t \qquad (5.6.2)$$

$$\Delta t'' = (1-\gamma)\Delta t \qquad (5.6.3)$$

Then the displacement as the wave reaches the external node is

$$\bar{U}^n = U^N \frac{(\Delta t'' + \Delta t)}{\Delta t} - U^{n-1} \frac{\Delta t''}{\Delta t} + \frac{\Delta t''(\Delta t + \Delta t'')\ddot{U}^n}{2} \tag{5.6.4}$$

and an approximation for the displacement of the external node at the end of the time step is

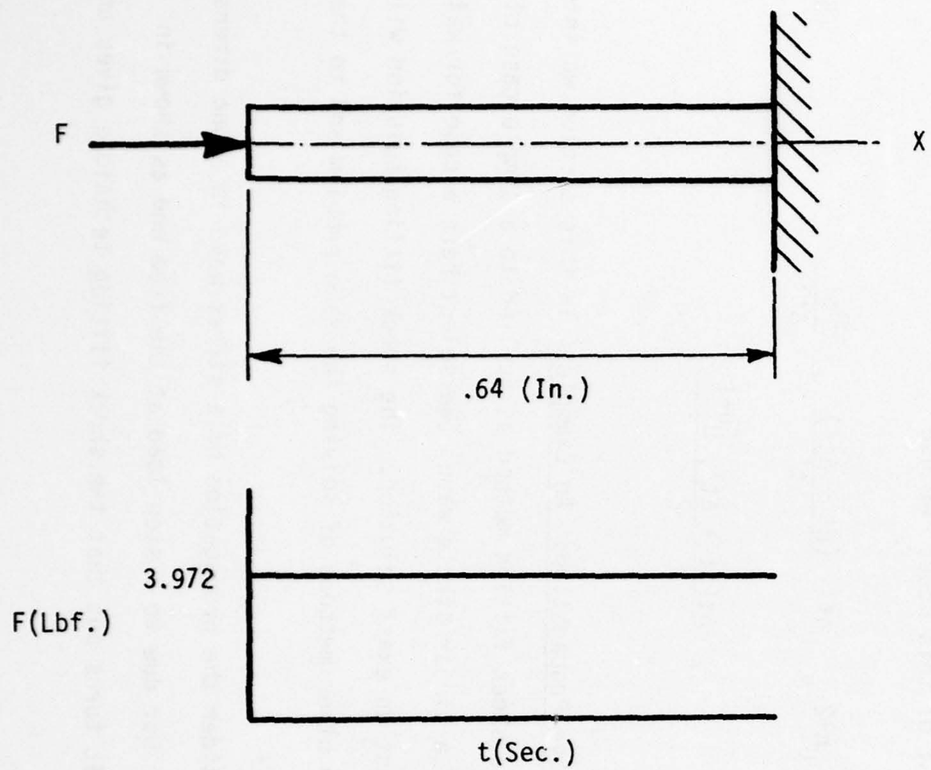
$$U^{n+1} = \bar{U}^n - V_n S^{n+1} \Delta t' \tag{5.6.5}$$

In the next time step we must again use a special procedure to compute the displacement of this node. We use

$$U^{n+2} = U^{n+1} \frac{(\Delta t + \Delta t')}{\Delta t'} - \frac{\bar{U}^n \Delta t}{\Delta t'} + \frac{\Delta t(\Delta t + \Delta t')\ \ddot{U}^{n+1}}{2} \tag{5.6.6}$$

V.7 Linear Wave Propagation: An Example. In this section we intend to demonstrate the shock fitting method as applied to a simple case of the propagation of a linear stress wave. We select here a case for which we can construct an exact solution. The shock fitting solution will be compared to other methods of solving the same problem and to the exact solution.

Consider the propagation of a stress wave in a one dimensional linear elastic bar due to a step load at the free end as shown in Figure 5.9. It turns out that the shock fitting technique gives an



Physical Parameters

Density = 1.0×10^{-4} $\frac{\text{Lbf.-Sec.}^2}{\text{In.}^4}$

Area = .0314 In.^2

Modulus of Elasticity = 63.25 Lbf./In.^2

Length = 0.64 In.

Figure 5.9 Physical Model for Propagation of a Linear Wave.

exact solution to this problem (including reflections). This is not too hard to understand since the piecewise linear finite element displacement model can exactly reproduce the constant strain state in front of and in back of the wave.

In Figure 5.10 the finite element shock fitting solution (Figure 5.10(a)) is compared to the parabolic regularization solution with $\beta = 0.8$ (Figure 5.10(b)) and the central difference solution (Figure 5.10(c)) at time point $t = .27 \times 10^{-4}$ sec. In Figure 5.11 the finite element shock fitting solution is shown at several subsequent time steps. The advantages of the method as applied to this problem are obvious

V.8 Growth of Shock Waves. In order to demonstrate the use of the shock fitting scheme, we consider the growth of a compression shock wave in a one-dimensional elastic bar of Mooney material. The constitutive equation for the Mooney material is given in (2.2.12) and (2.2.18). The physical parameters of the problem as well as the forcing function are shown in Figure 5.12.

In the calculations to be presented here the implicit shock position method was used with the linear trial functions (the DIS 1 set). The same equations as described in V.5 and V.6 were used in the calculations with one exception. In the standard elements an explicit artificial viscosity term of the following form was added to the equations for each element:

$$\Delta t^{0.8} \int_{I_e} \left. \frac{\partial \sigma(U_X)}{\partial U_X} \right|_{U=U^n} \frac{U_X^n - U_X^{n-1}}{\Delta t} \psi_{\beta,X} dX \quad \beta=1,2$$

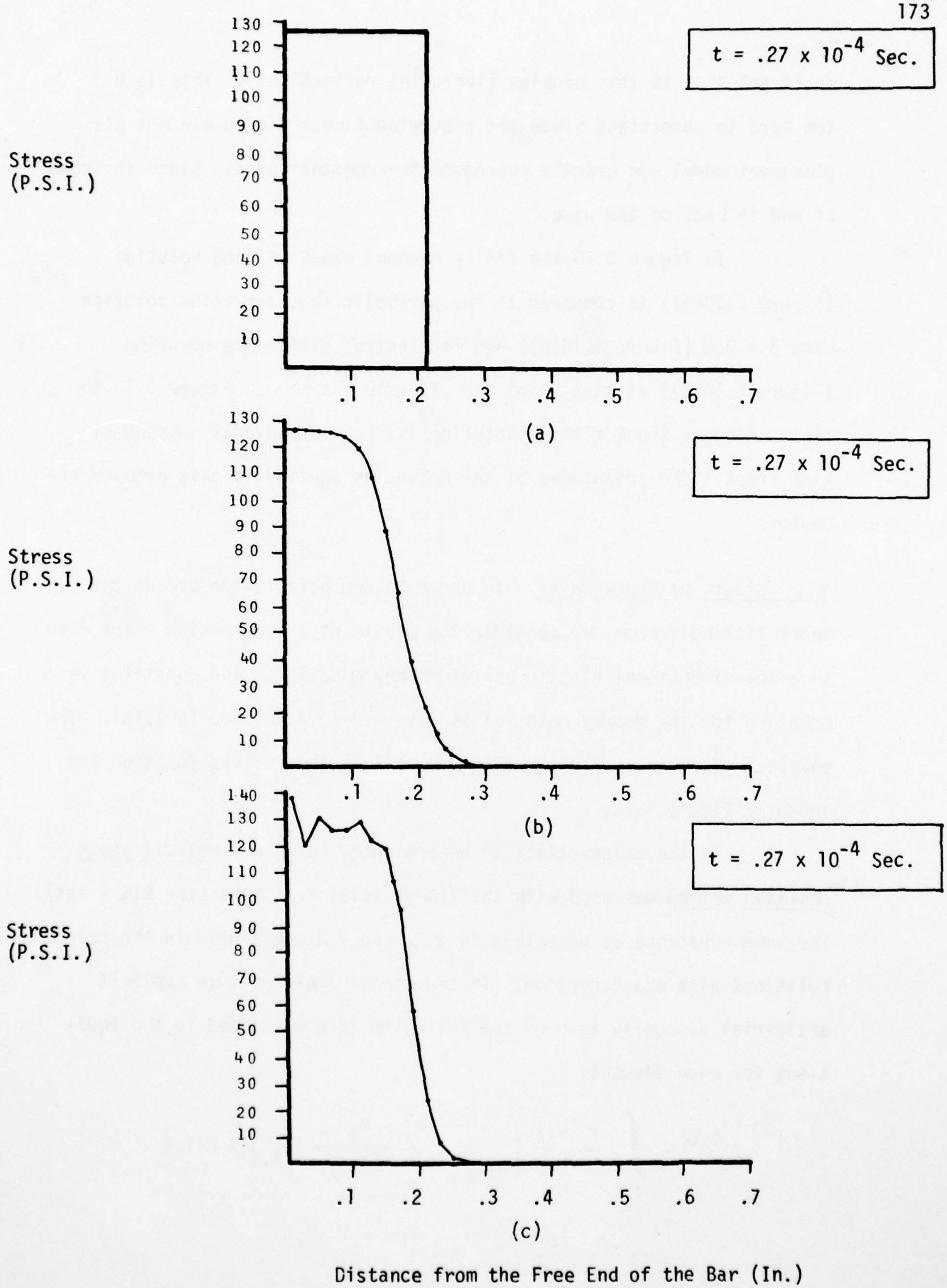
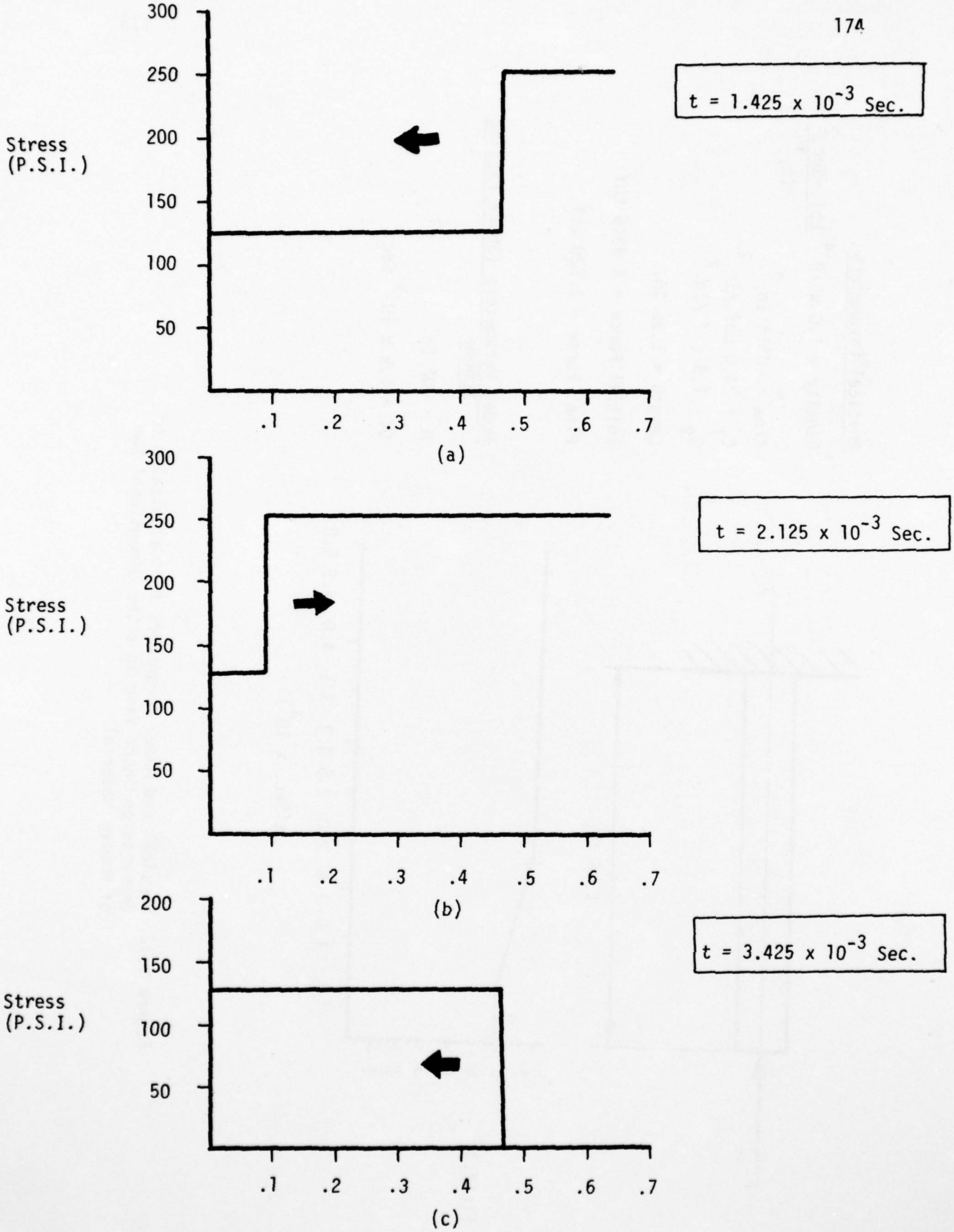
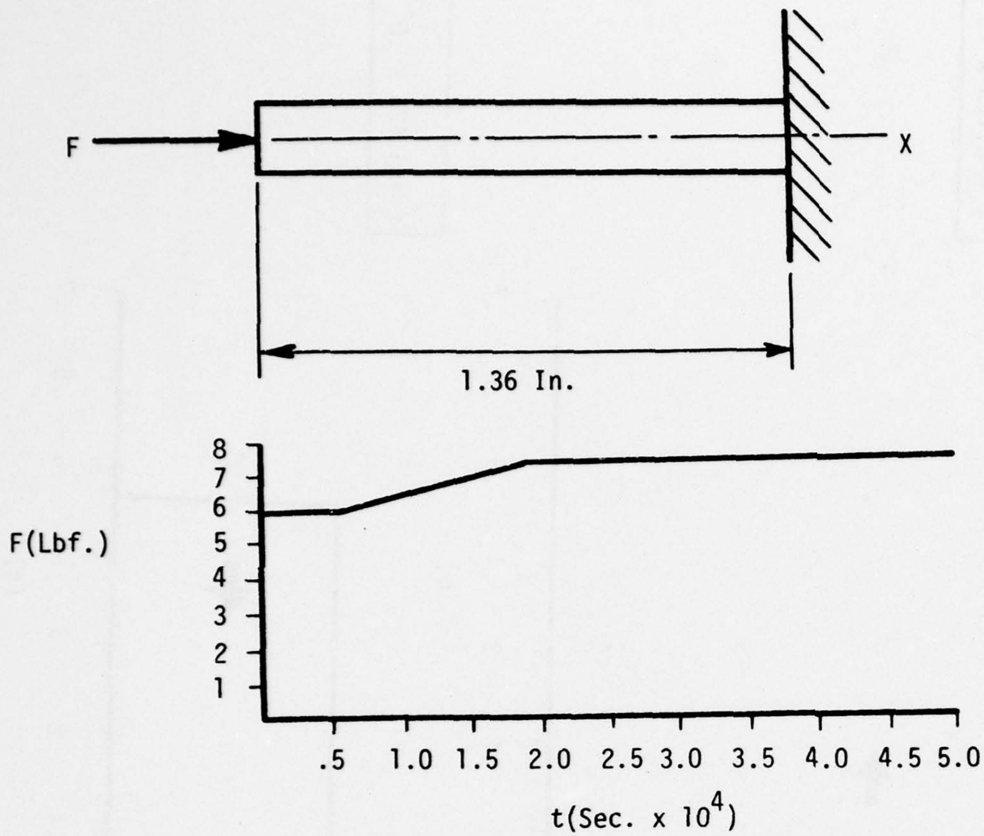


Figure 5.10 Linear Wave Propagation Calculation by Several Methods



Distance From the Free End of the Bar (In.)

Figure 5.11 Linear Wave Propagation by the Shock Fitting Method



Physical Parameters

Density = 1.0×10^{-4} $\frac{\text{Lbf.-Sec.}^2}{\text{In.}^4}$

Area = .0314 In.^2

C_1 = 24.0 Lbf./In.^2

C_2 = 1.5 Lbf./In.^2

Length = 1.36 In.

Initial Force = 5.9346 Lbf.

Final Force = 7.220 Lbf.

Model Parameters (Shock Fitting Scheme

h = .02 In.

DT = 1.5×10^{-7} Sec.

Figure 5.12 Physical and Model Parameters for Calculation of Compression Shock Wave in a One-Dimensional Bar of Mooney Material

We provide the following explanation for the use of this artificial viscosity: Since the applied load shown in Figure 5.12 has a discontinuity in slope, a jump in acceleration is introduced in the rod at the discontinuity. Thus an acceleration wave is formed in the rod. From the results of Chapter IV (in particular Theorem 4.2) it is clear that the standard Galerkin technique will not converge to acceleration wave solutions. The only choice to insure convergence is to regularize the approximation or to develop a fitting technique for acceleration waves. To regularize the approximation we usually add artificial damping. We have chosen here to regularize the acceleration wave rather than treat it by some fitting technique for acceleration waves. Essentially since our primary goal here is to model the shock wave, we smeared the acceleration wave rather than treat it through some fitting technique for acceleration waves. No artificial viscosity was used in the special element at the shock wave.

In Figure 5.13 the variation of the shock strength S is given as a function of time. It is compared to the theoretical upper bound of the shock strength which would be attained if there was no dissipation of the strength of shock in the shock propagation process. In Figure 5.14 the variation of the intrinsic speed of the shock wave is plotted as a function of time. We note here one of the advantages of the finite element/shock fitting scheme. The shock strength and intrinsic wave speed are extremely accessible in that they are dependent variables in the finite element model. The variation or oscillation in the intrinsic wave speed seen in Figure 5.14 is due to inertial coupling across the wave induced by the nondiagonal mass matrix.

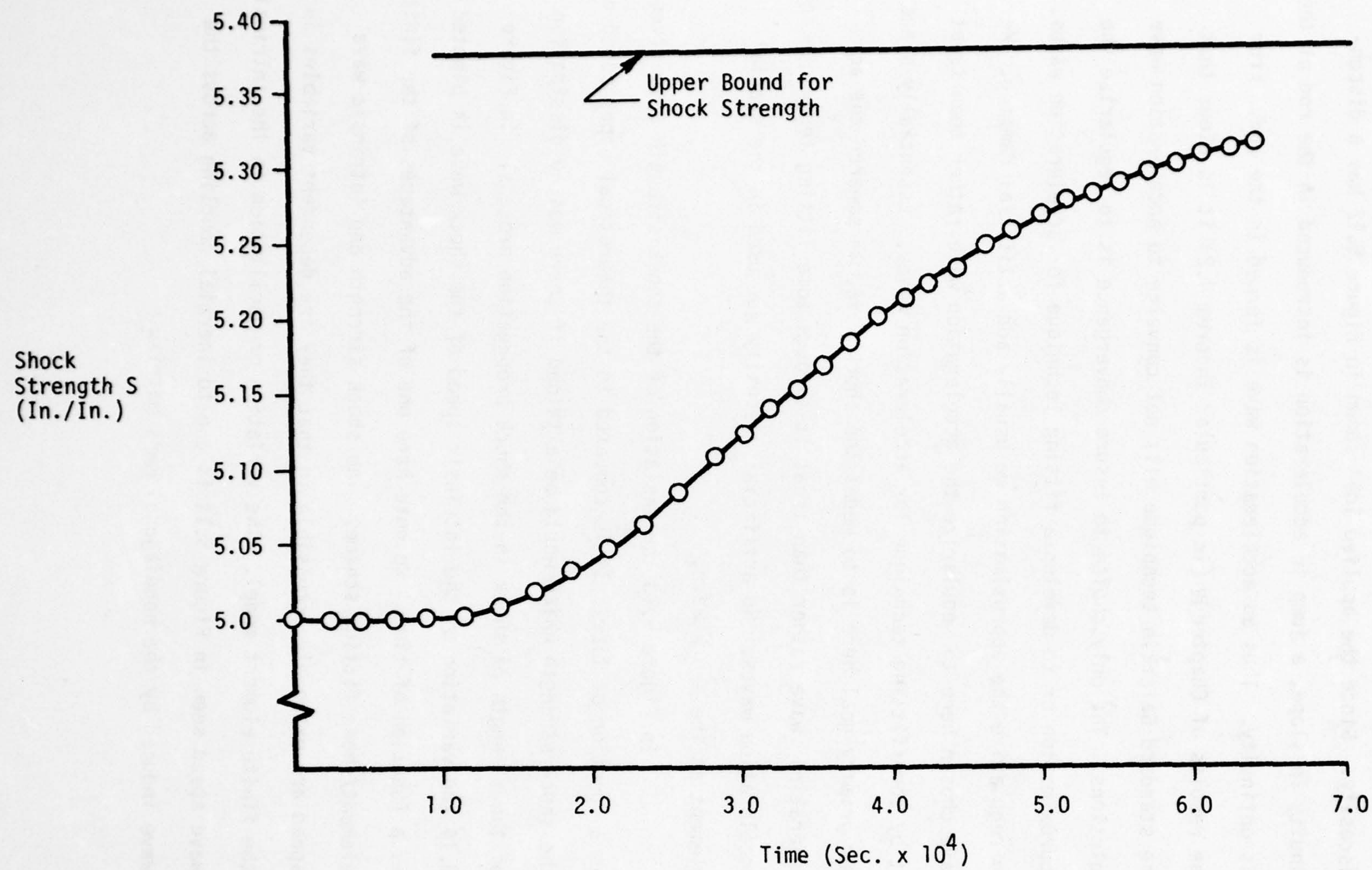


Figure 5.13 Growth of Shock Strength for the Compression Shock Computed by the Shock Fitting Scheme.

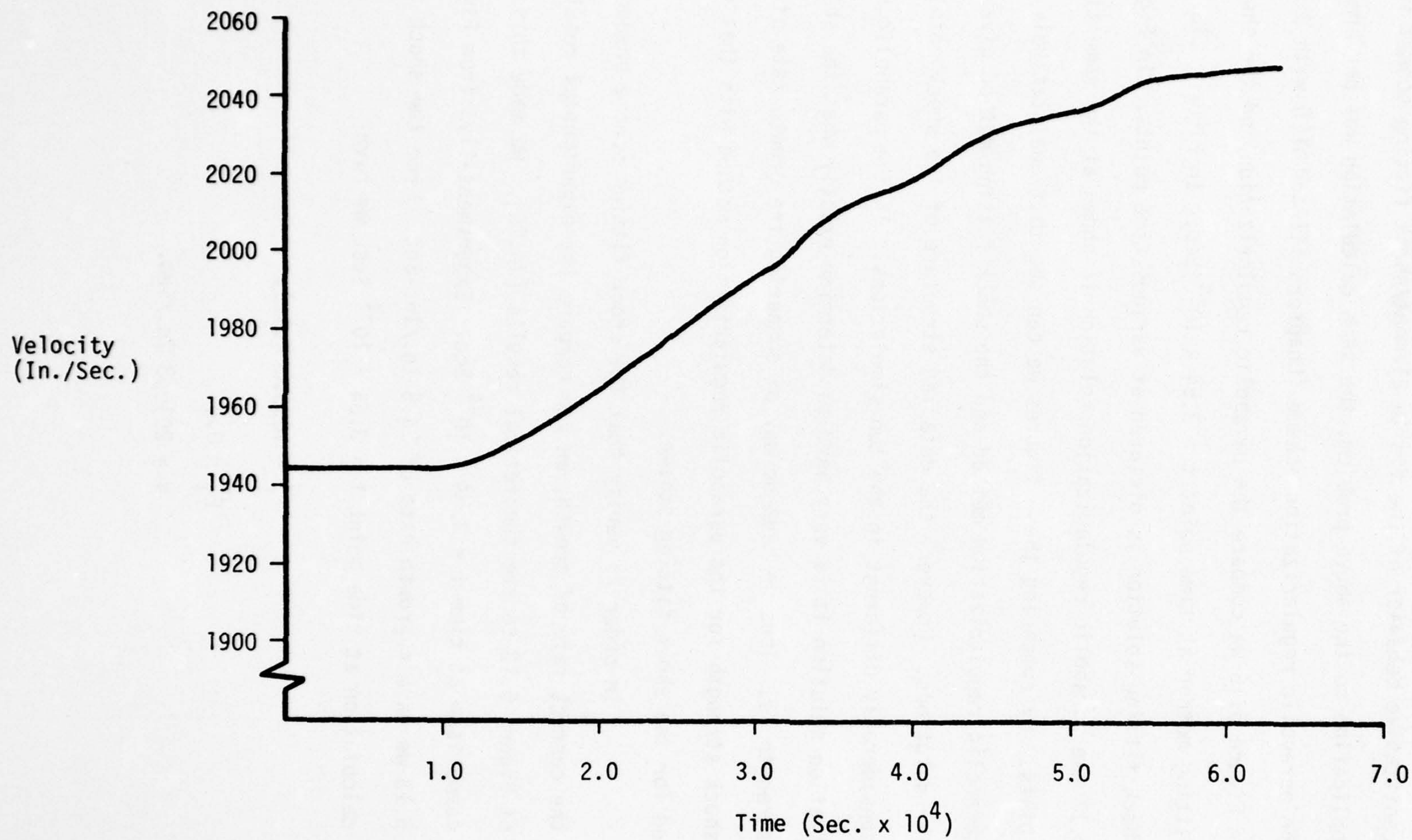


Figure 5.14 Intrinsic Velocity of the Compression Shock Wave Computed by the Shock Fitting Scheme.

To provide some basis for evaluating the qualitative and quantitative behavior of the finite element/shock fitting scheme in application to the above problem, the same calculation was performed with the parabolic regularization scheme (Chapters VII and VIII) with $\beta = 0.8$. In Figure 5.15 we compare the parabolic regularization and the shock fitting method at time point $t = 3.54 \times 10^{-4}$ Sec. In Figure .16 the shock fitting solution is pictured at various time points. In Figure 5.17 the parabolic regularization solution is shown at the same time points. By comparing these figures we can see that qualitatively the parabolic regularization method and the shock fitting method give similar solutions. However, the detailed structure of the shock point is considerably different in the two simulations. In the parabolic regularization solution it is very hard to determine exactly what the shock strength is. Thus we have no way of comparing the growth rate of the shock strength for the parabolic regularization method with that obtained for the shock fitting scheme.

In order to verify that the shock fitting scheme produces the correct rate of growth, we can compare the experimental results of Figure 5.13 to the theoretical results [74,75]. We made this comparison at time $t = 3.54 \times 10^{-4}$ Sec. Experimentally, from Figure 5.13 we have a growth rate of 94.5 In./In.-Sec. From the shock fitting calculation at time point $t = 3.54 \times 10^{-4}$ Sec. we have

$$E_T^- = 1093.0 \text{ P.S.I.}$$

$$E_T^+ = 0.0$$

$$V = 2010.3 \text{ In./Sec.}$$

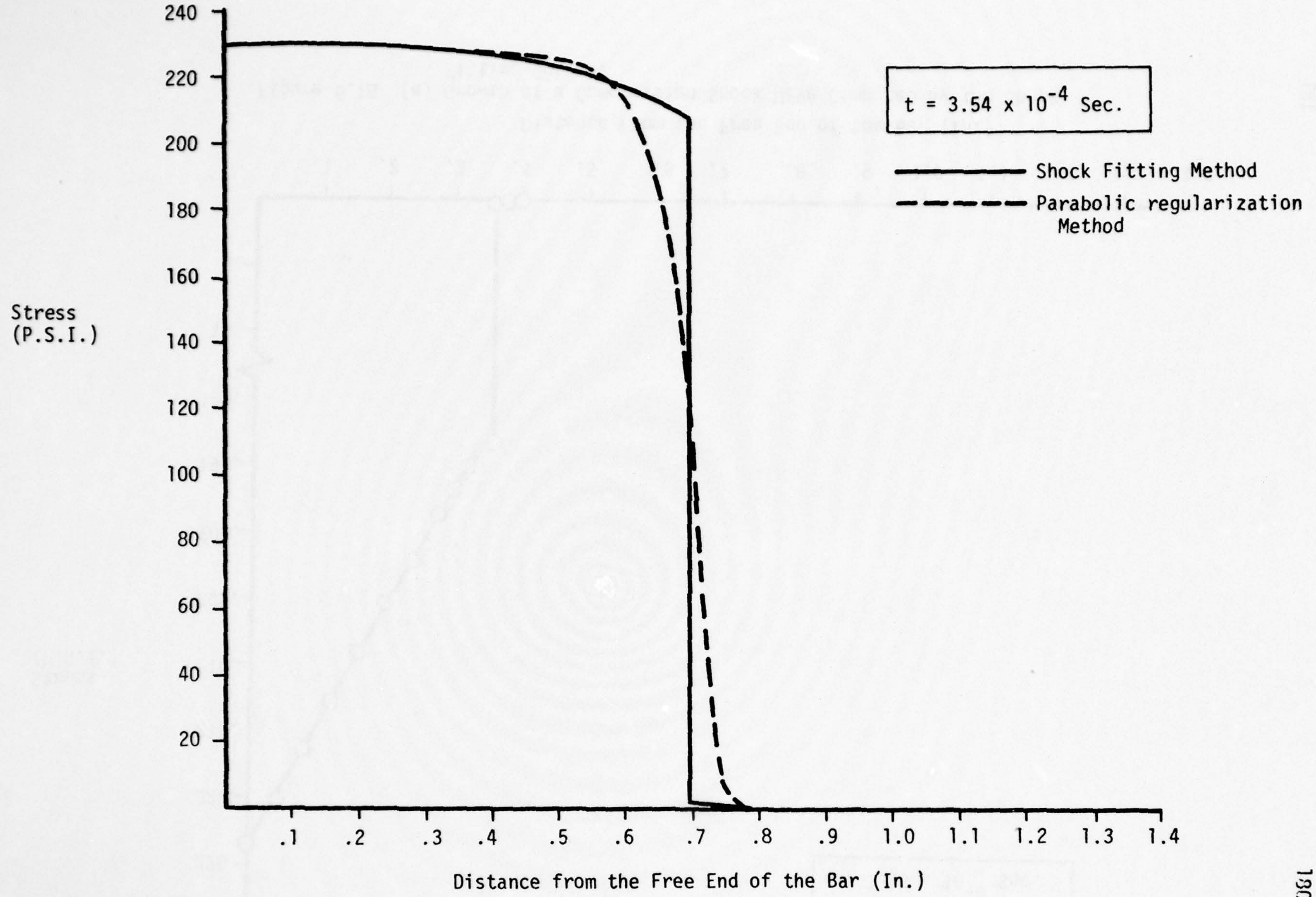


Figure 5.15 A Comparison of the Shock Fitting Solution and the Parabolic Regularization Solution at $t = 3.54 \times 10^{-4}$ Sec.

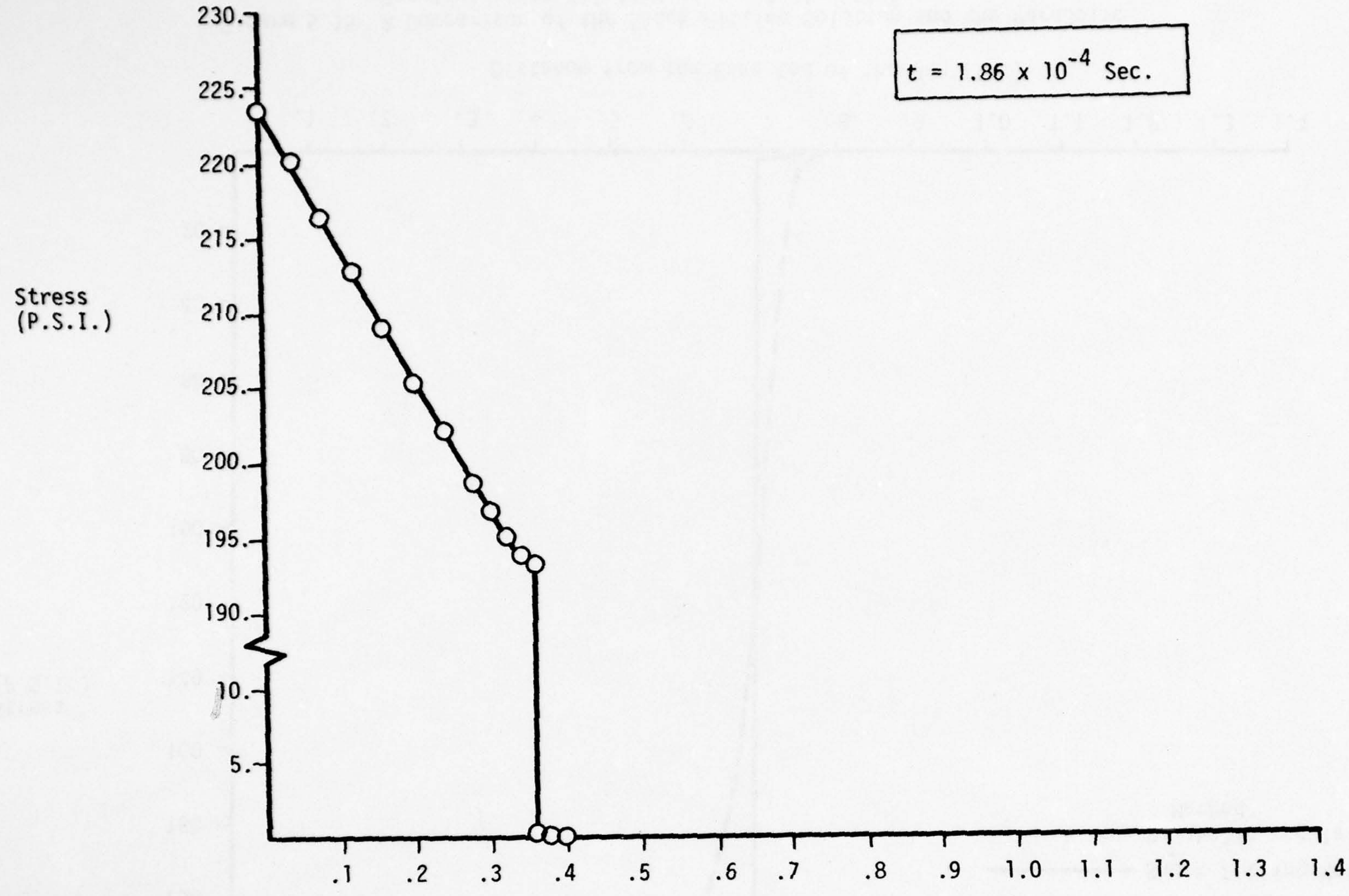


Figure 5.16 (a) Growth of a Compression Shock Wave Computed by the Shock Fitting Scheme.

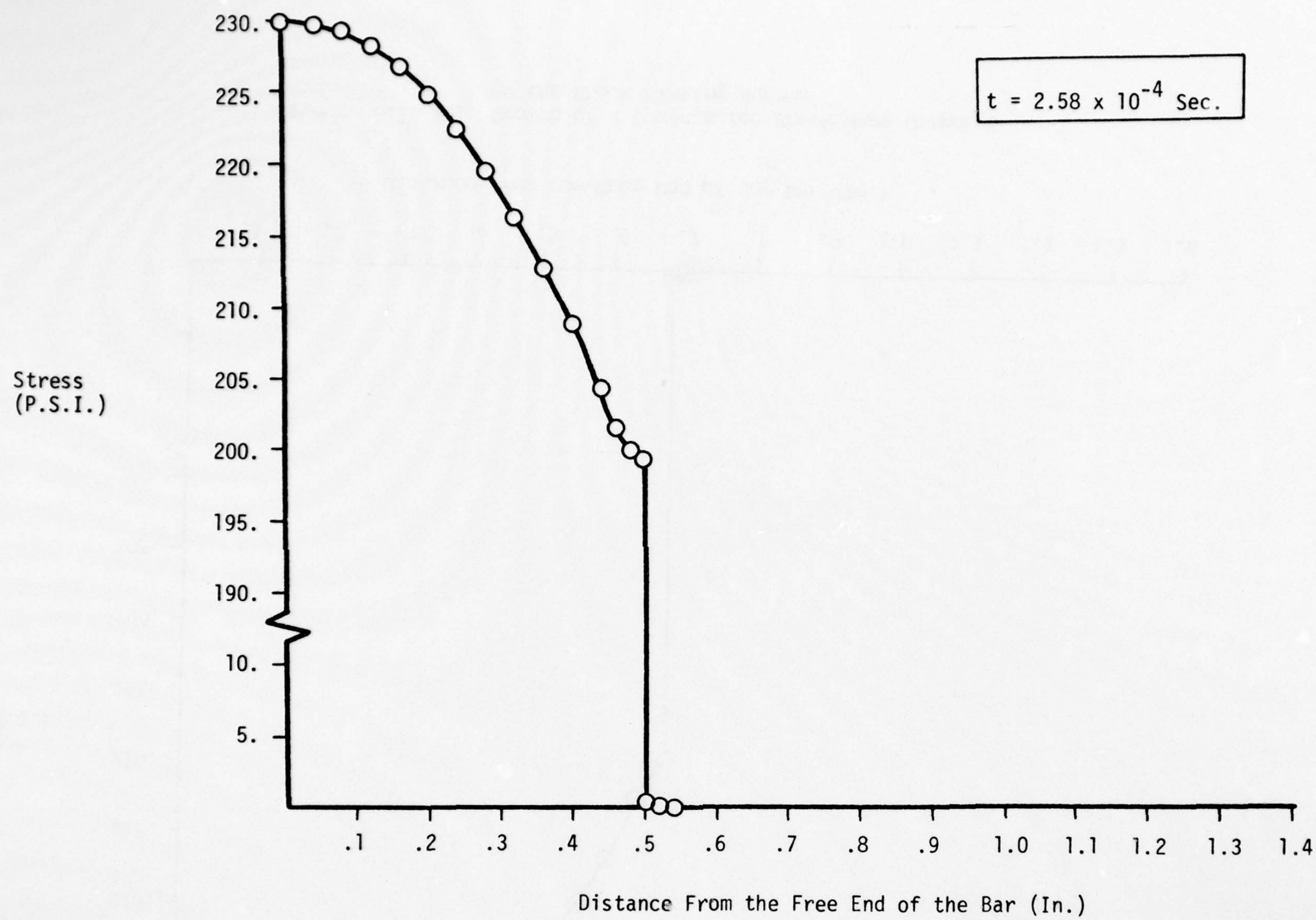


Figure 5.16 (b) Growth of a Compression Shock Wave Computed by the Shock Fitting Scheme.

AD-A031 238 TEXAS UNIV AT AUSTIN DEPT OF AEROSPACE ENGINEERING AN--ETC F/G 20/11
A STUDY OF CONVERGENCE AND STABILITY OF FINITE ELEMENT APPROXIM--ETC(U)
AUG 76 J T ODEN, L C WELLFORD, C T REDDY DAAG29-76-G-0022
UNCLASSIFIED ARO-11860.9-M NL

3 OF 4
ADA031238

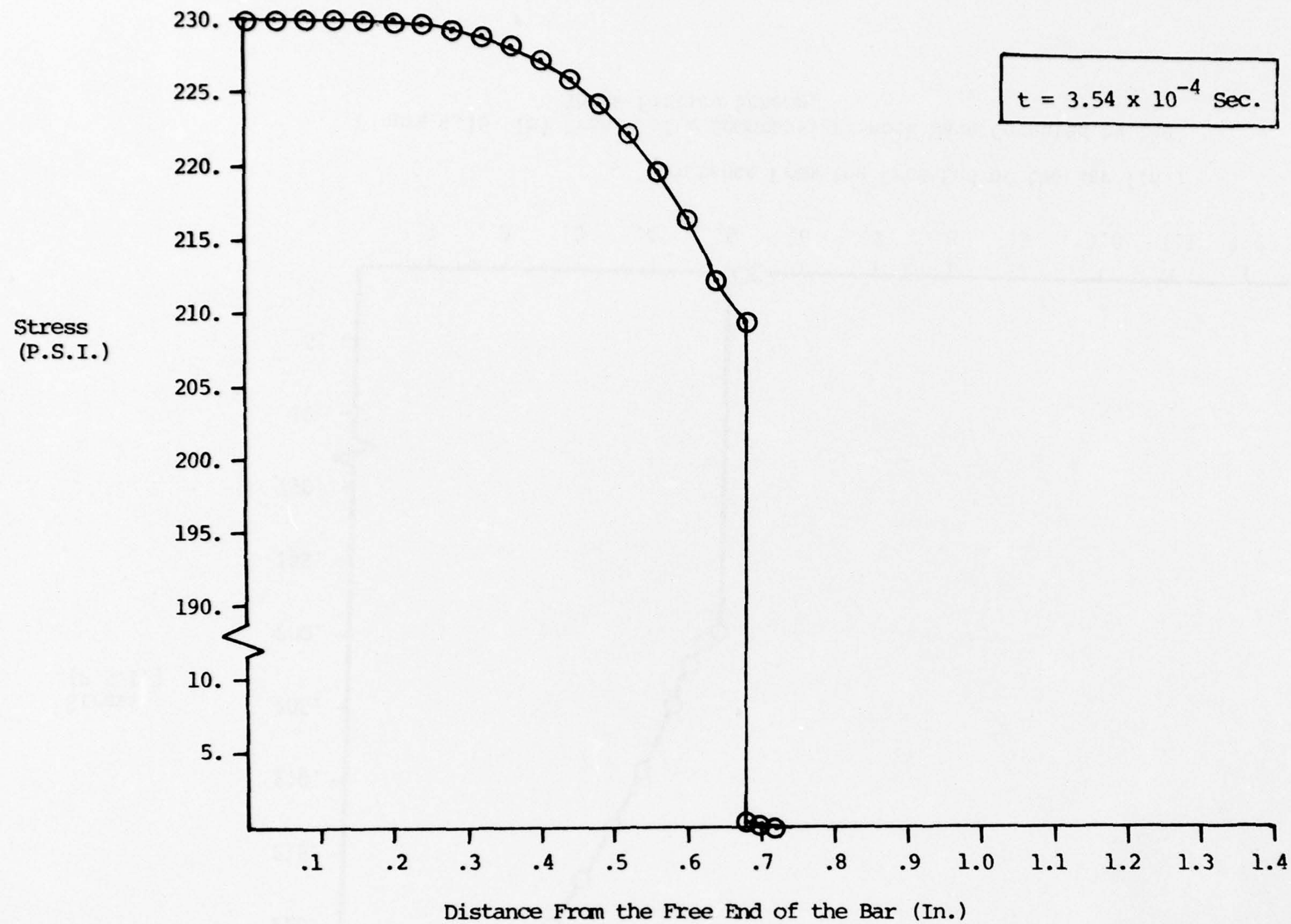


Figure 5.16 (c) Growth of a Compression Shock Wave Computed by the Shock Fitting Scheme

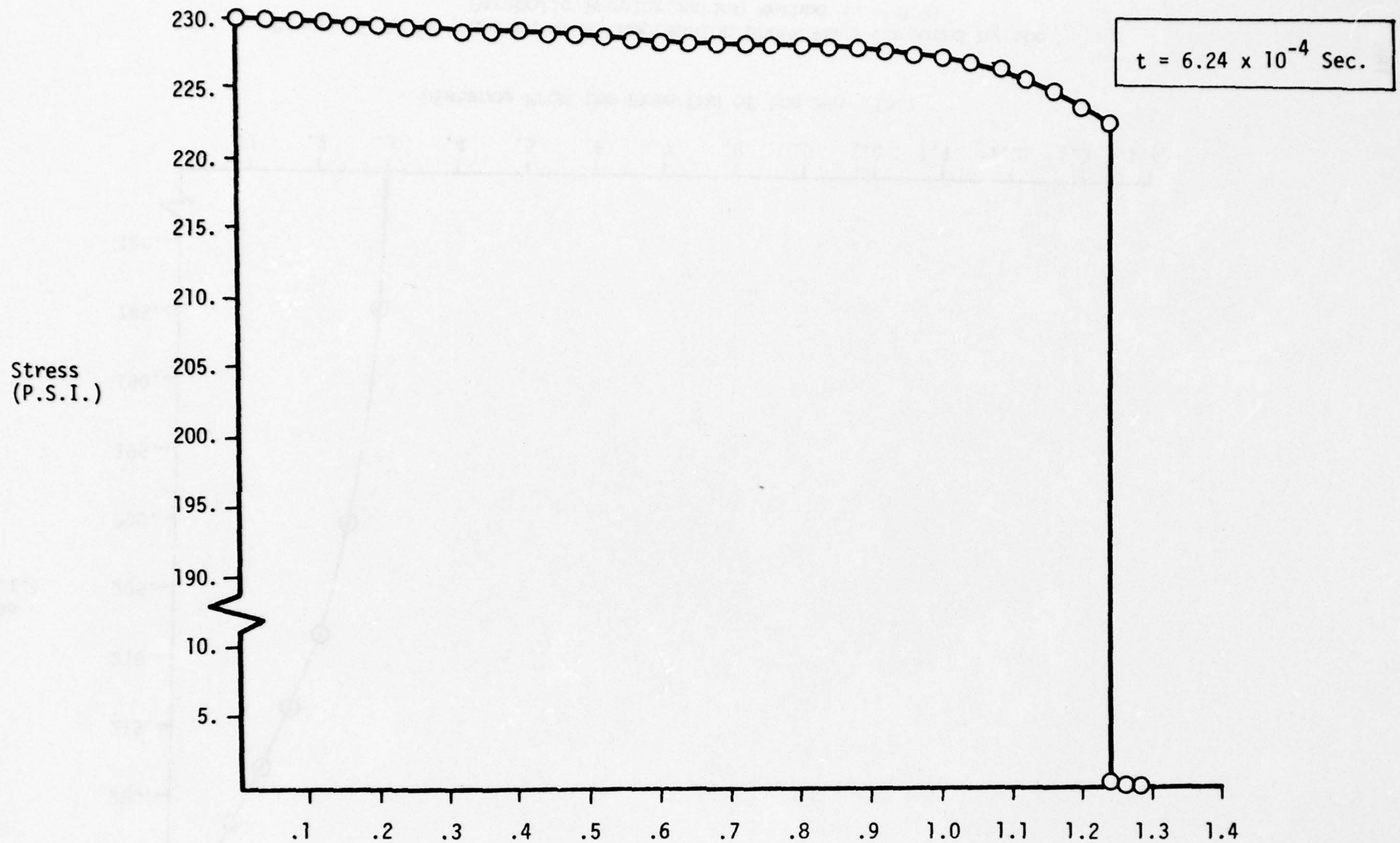


Figure 5.16 (d) Growth of a Compression Shock Wave Computed by the Shock Fitting Scheme.

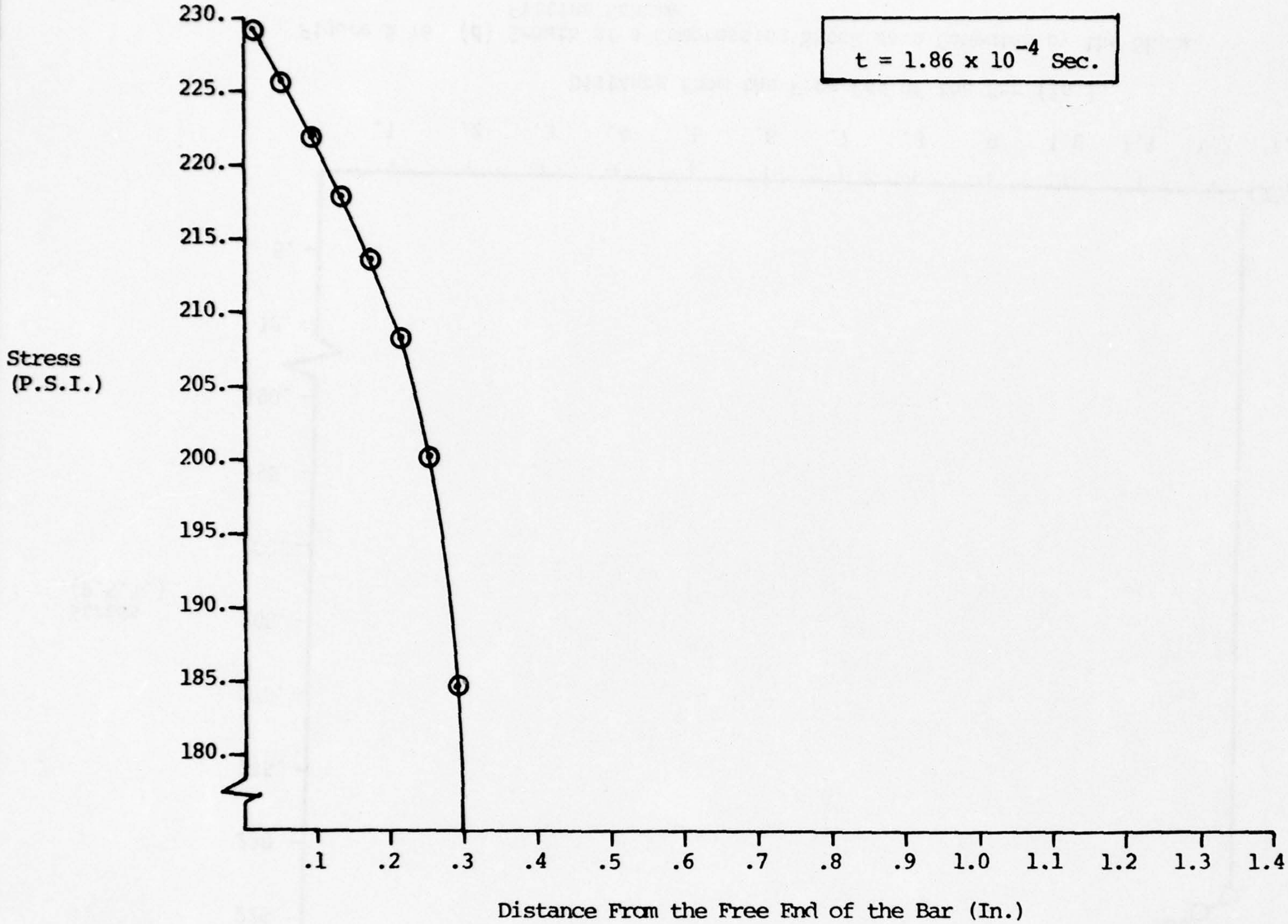


Figure 5. 17 (a) Growth of a Compression Shock Wave Computed by the Parabolic Regularization Method ($\alpha = 0.8$)

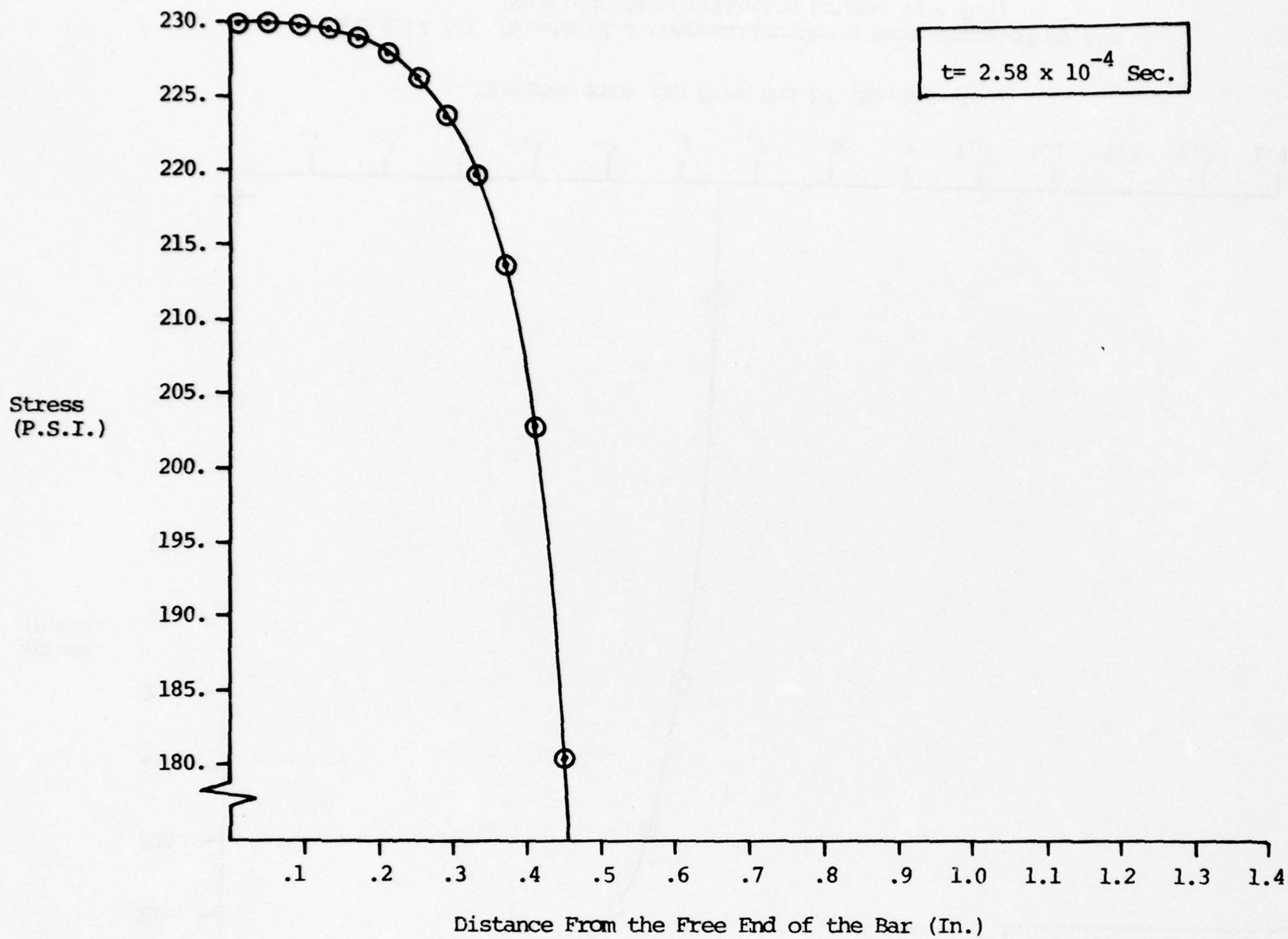


Figure 5.17 (b) Growth of a Compression Shock Wave Computed by the Parabolic Regularization Scheme ($\alpha = 0.8$)

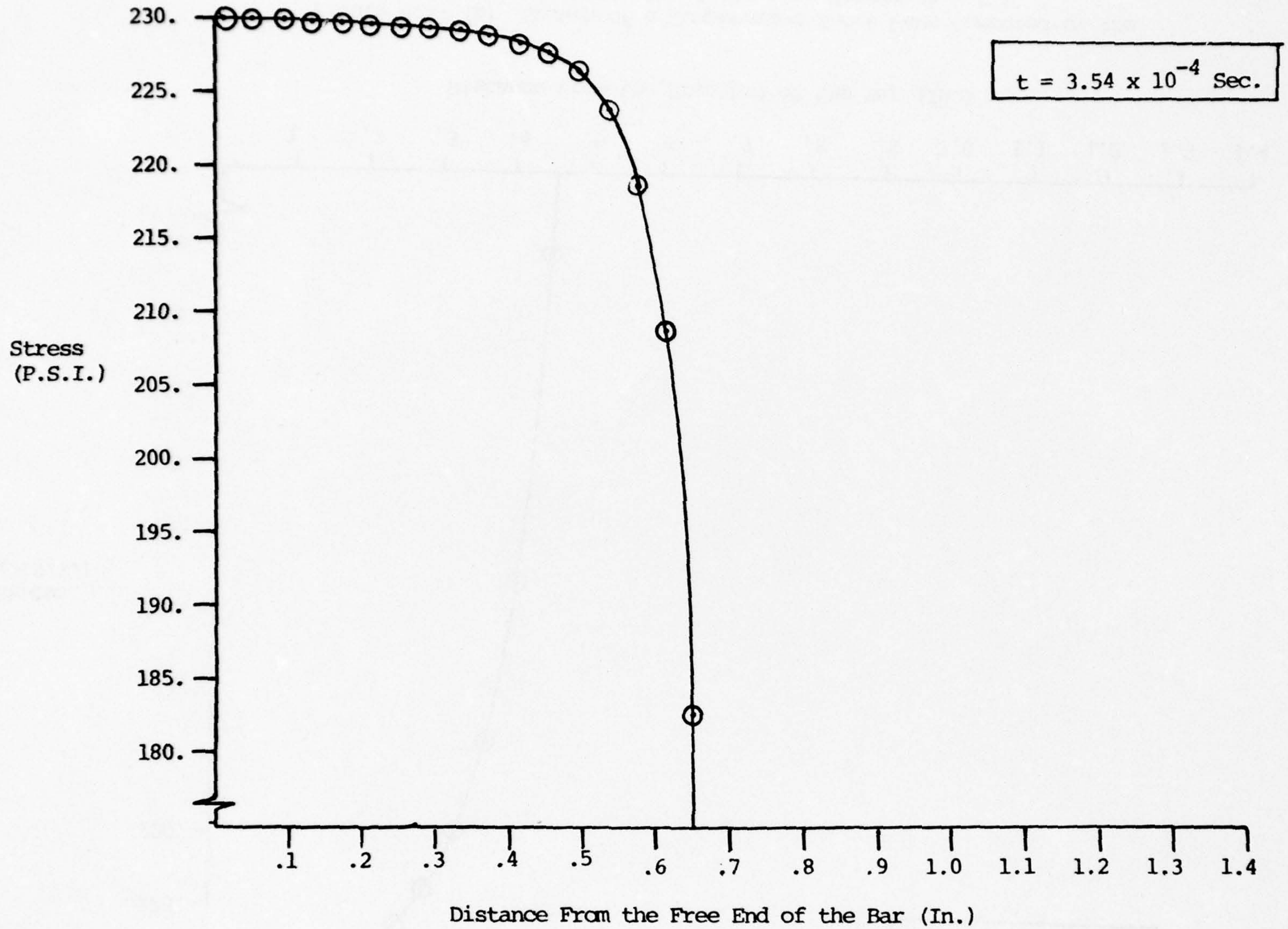


Figure 5.17 (c) Growth of a Compression Shock Wave Computed by the Parabolic Regularization Scheme ($\alpha = 0.8$)

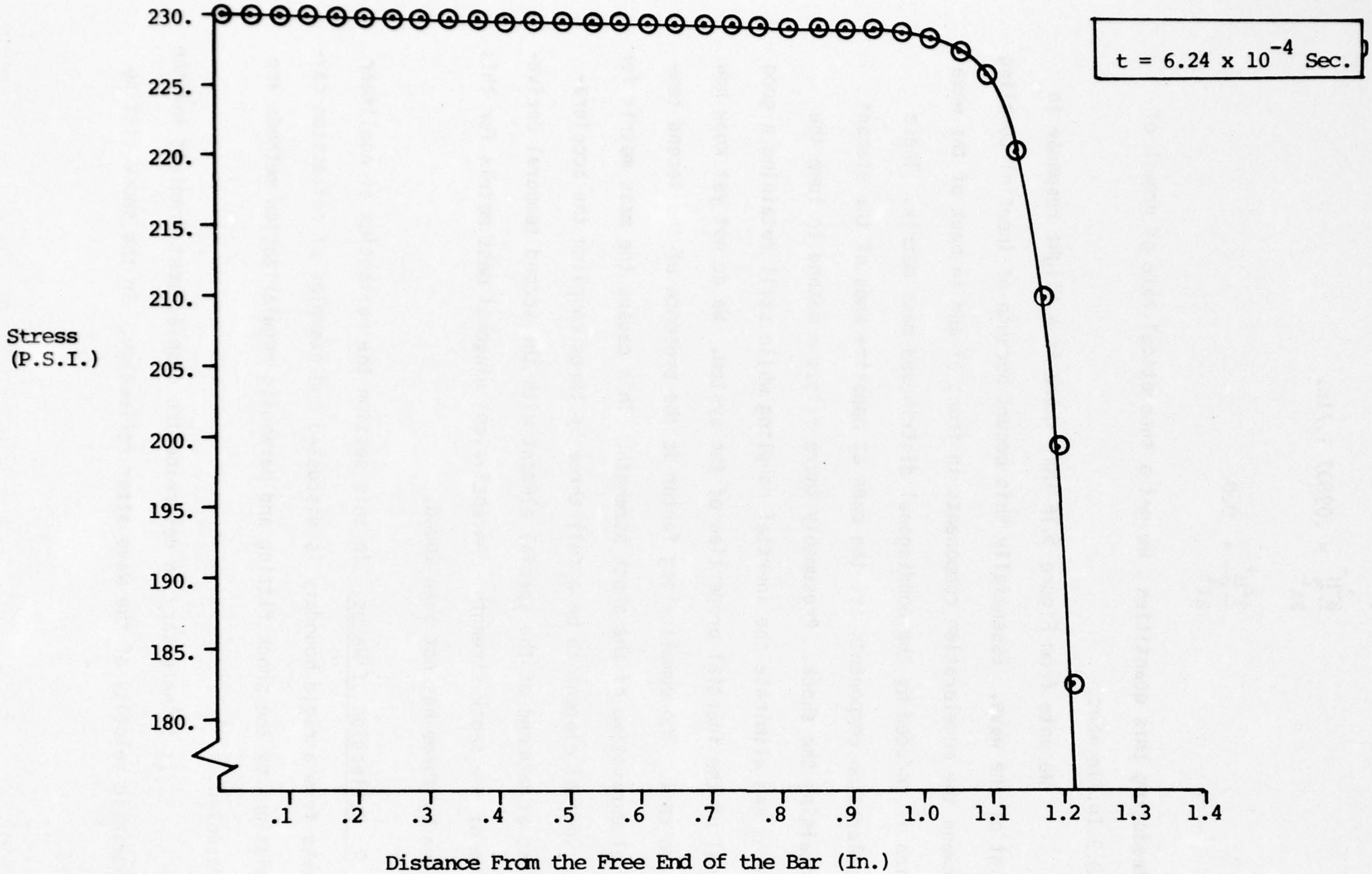


Figure 5.17 (d) Growth of a Compression Shock Wave Computed by the Parabolic Regularization Method ($\alpha = 0.8$)

$$\frac{\partial^2 U^-}{\partial X^2} = .09837 \text{ 1./In.}$$

$$\frac{\partial^2 U^+}{\partial X^2} = 0.0$$

Introducing this quantities , we get a theoretical rate of growth of 118.3 In./In.-Sec.

We note from Figure 5.16 that there is a slight response in front of the wave. Essentially this occurs because of inertial coupling between the acceleration components in front of and in back of the wave which is induced by the nondiagonal distributed mass matrix. These acceleration components are the ones at opposite ends of the element containing the shock. Presumably there exists a method to lump the masses and eliminate the inertial coupling while still retaining a good model of the inertial properties of the system. We do not yet know how to do this. The complicating factor is the presence of the second temporal derivative of the shock strength. This causes the mass matrix for the special element to be a full three by three coupling the accelerations at each end of the special element with the second temporal derivative of the shock strength. An equivalent diagonal mass matrix for this three by three has not been found.

V .9 Reflection of Waves. In this section the reflection of nonlinear waves from a rigid boundary is discussed and examples of reflection calculations by the shock fitting and parabolic regularization methods are presented.

It is necessary to determine the displacement gradient and the intrinsic velocity of the wave after reflection. In the shock fitting

method we require that the momentum jump condition and the kinematical compatibility condition be satisfied. Noting that after reflection the particle velocity at the wall must be zero, these conditions imply that

$$-\rho \frac{dU^-}{dX} V^2 + \rho \frac{dU^+}{dX} V^2 - \sigma^+ + \sigma^- = 0$$

$$V \frac{dU^-}{dX} - V \frac{dU^+}{dX} + \dot{U}^- = 0 \tag{5.9.1}$$

dU^-/dX, σ^-, and $\dot{U}^-$ are known properties of the incoming wave. dU^+/dX and V are the unknowns of the problem.

Equation (5.9.1) can be solved by the Newton-Raphson method. Let $\underset{\sim}{X} = \{V, dU^+/dX\}$. Then let

$$\underset{\sim}{F} = \left\{ \begin{array}{l} -\rho \frac{dU^-}{dX} X_1^2 + \rho X_2 X_1^2 - \sigma^+ + \sigma^- \\ X_1 \frac{dU^-}{dX} - X_1 X_2 + \dot{U}^- \end{array} \right\} \tag{5.9.2}$$

Then as an iterative procedure we require that

$$\underset{\sim}{X}^{n+1} = \underset{\sim}{X}^n - (\underset{\sim}{J}^n)^{-1} \underset{\sim}{F}^n \tag{5.9.3}$$

where $\underset{\sim}{J}$ is the Jacobian matrix defined by

$$\underset{\sim}{J} = [\frac{\partial F_i}{\partial X_j}] \tag{5.9.4}$$

If we consider the Mooney material (defined by equations (2.2.12) and (2.2.18)), the following explicit expressions for the entries in $\underset{\sim}{J}$ can be developed:

$$J_{11} = -2\rho\frac{dU^-}{dX}X_1 + 2\rho X_2X_1$$

$$J_{12} = \rho X_1^2 - 2C_1 - \frac{4C_1}{(1+X_2)^3} - \frac{6C_2}{(1+X_2)^4}$$

$$J_{21} = \frac{dU^-}{dX} - X_2 \tag{5.9.5}$$

$$J_{22} = -X_1$$

We have obtained very good results by using the reflected wave for a linear material as an initial guess to start the interation. Then we set

$$X_1^o = -\overline{V} \tag{5.9.6}$$

$$X_2^o = \frac{dU^-}{dX} + S$$

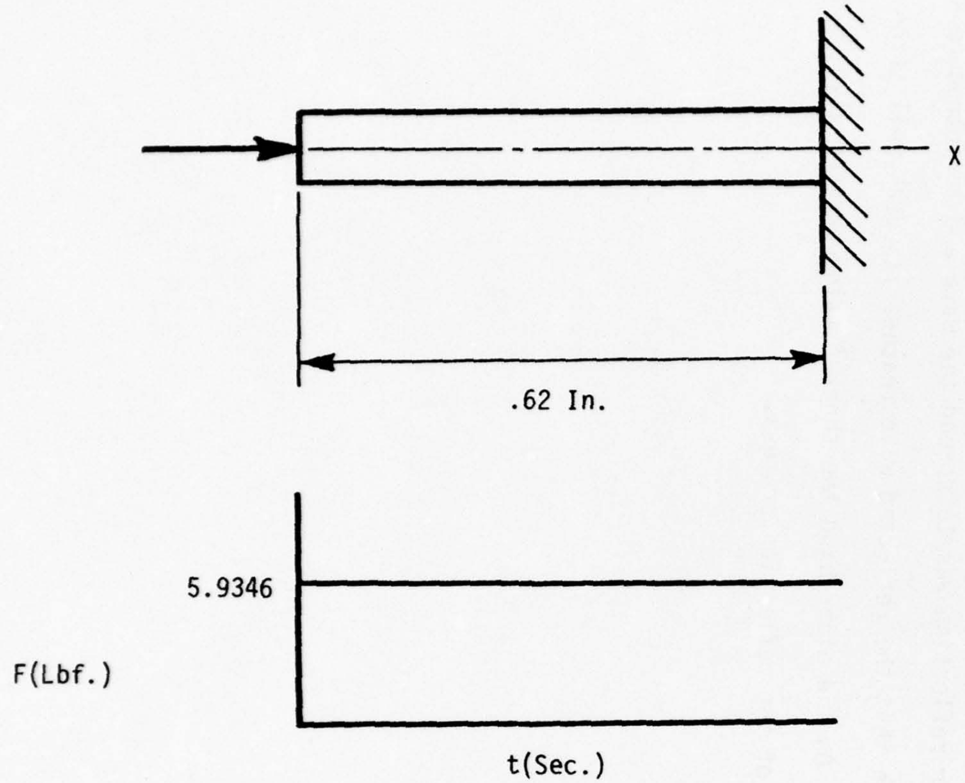
where $\overline{V}$ is the intrinsic velocity of the wave prior to reflection and S is the shock strength. Convergence occurs in at most 7 steps of the procedure (5.9.3). As a final step in defining the reflected wave, we set

$$S = \frac{dU^-}{dX} - X_2 \tag{5.9.7}$$

and proceed with the calculations.

Of course in the reflection calculation by shock smearing methods the reflection is "automatically accounted for." The reflection is a smooth transition as compared to the reflection induced by the scheme (5.9.3) for the shock fitting scheme which is abrupt. We believe that this smoothing of the reflection in the parabolic regularization method and other shock smearing method is a serious drawback and pollutes the approximation with error after the reflection process is completed.

We consider the reflection of a nonlinear wave in a one-dimensional rod of Mooney material. The physical parameters are presented in Figure 5.18. In Figure 5.19 the shock fitting solutions is presented before and after reflection. In Figure 5.20 we present the same problem except that the calculations are performed with the parabolic regularization method with $\alpha = 0.8$. Figure 5.20(a) and 5.20(b) correspond to the time points presented in Figure 5.19 for the shock fitting calculation. At time point $t = 3.1 \times 10^{-4}$ Sec. the shock fitting reflected wave is fully developed (650.4 P.S.I.); however, the parabolic regularization reflected wave is just beginning to develop. In Figure 5.20(c) at $t = 3.6 \times 10^{-4}$ Sec. the parabolic regularization reflected wave is fully developed (625.0 P.S.I.). The effect of the parabolic regularization reflection was to spread the wave and to introduce a phase error in that the reflected wave reaches its peak well after the actual wave. On the other hand the shock fitting program gives a detailed model of the reflection process.



<u>Physical Parameters</u>

Density = 1.0×10^{-4} $\dfrac{\text{Lbf.-Sec.}^2}{\text{In.}^4}$

Area = .0314 In.^2

C_1 = 24.0 Lbf./In.2

C_2 = 1.5 Lbf./In.^2

Length = .62 In.

Figure 5.18 Physical Parameters for the Calculation of Reflection of a Nonlinear Wave.

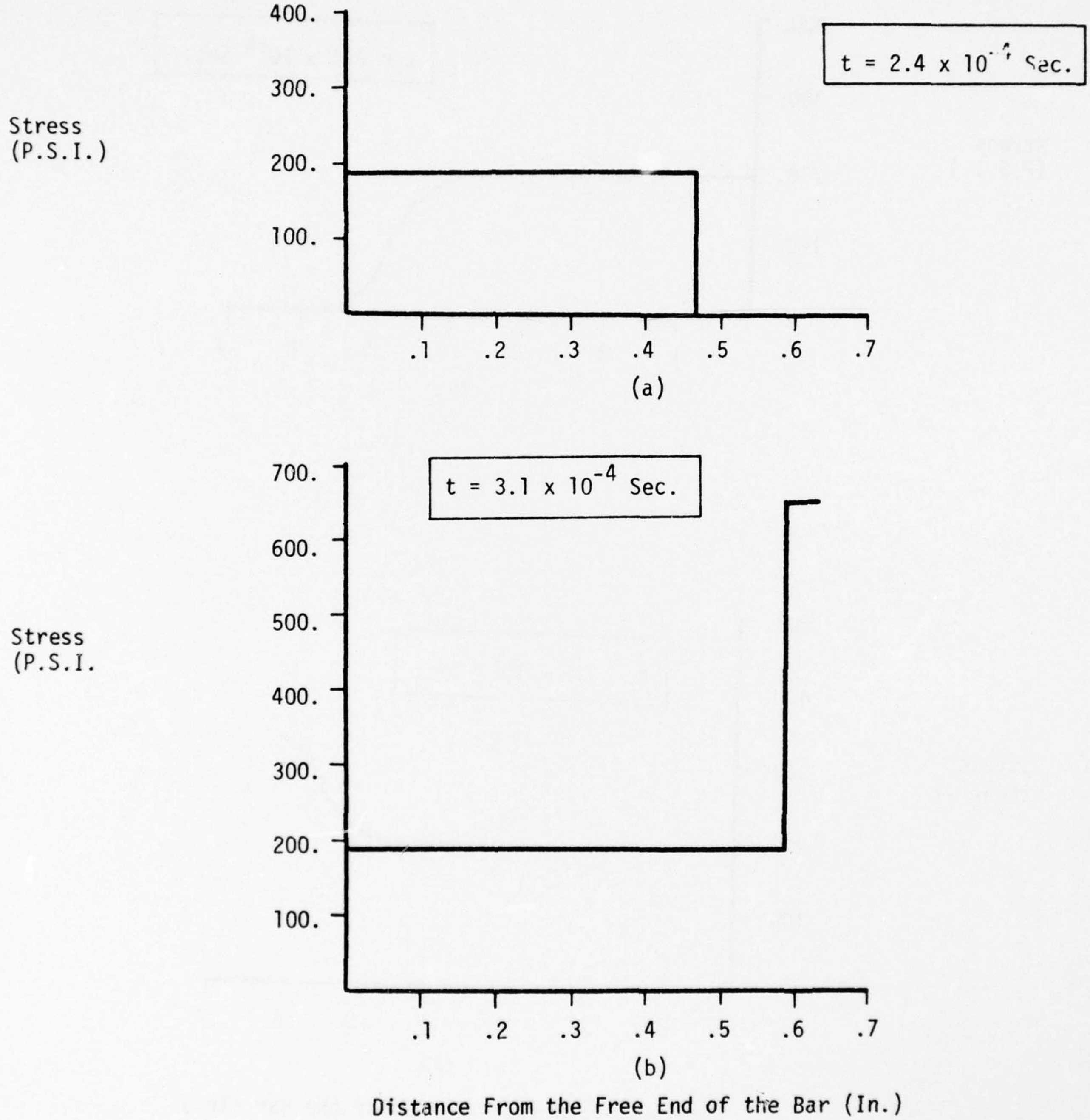


Figure 5.19 Reflection of a Nonlinear Wave by the Shock Fitting Method.

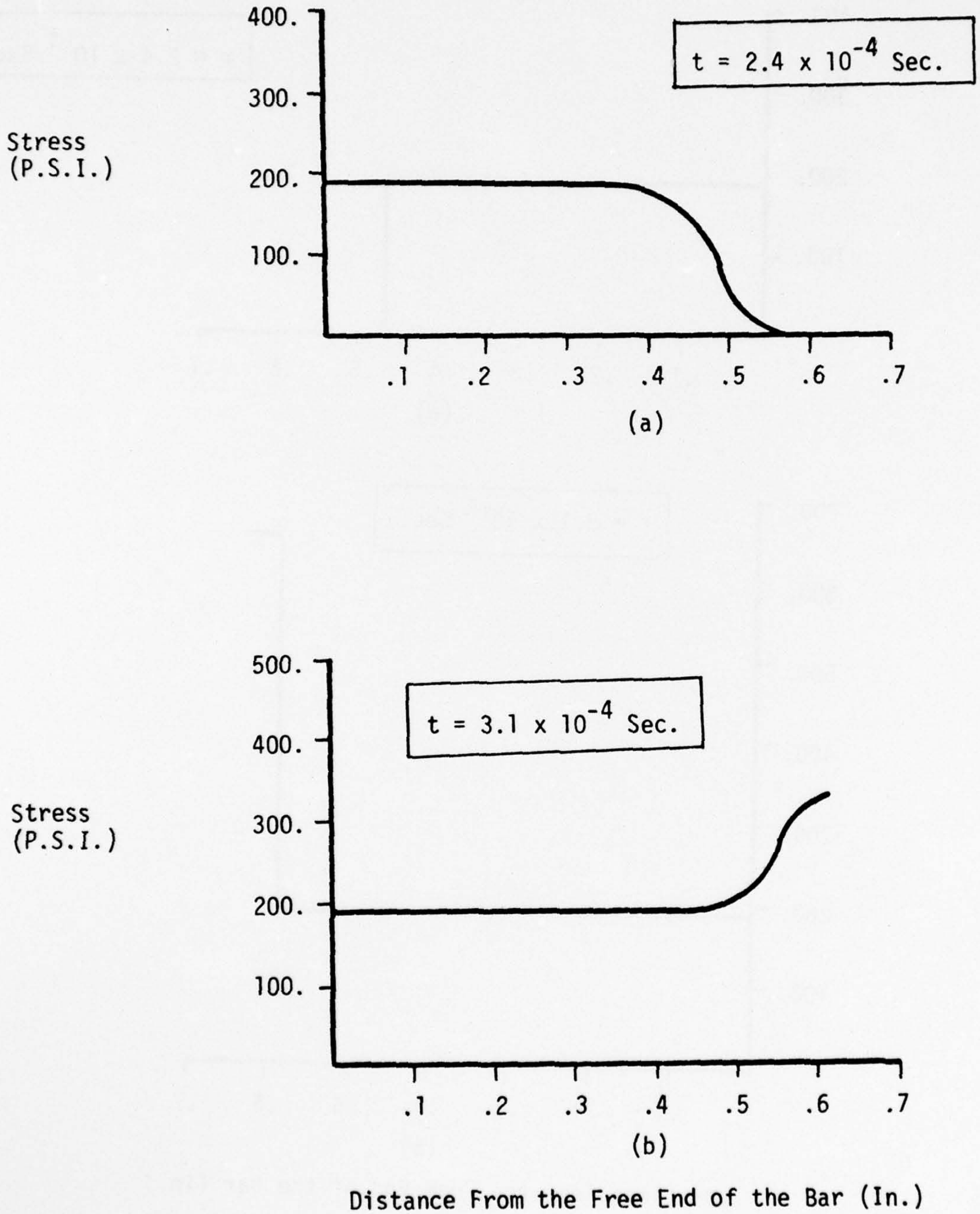
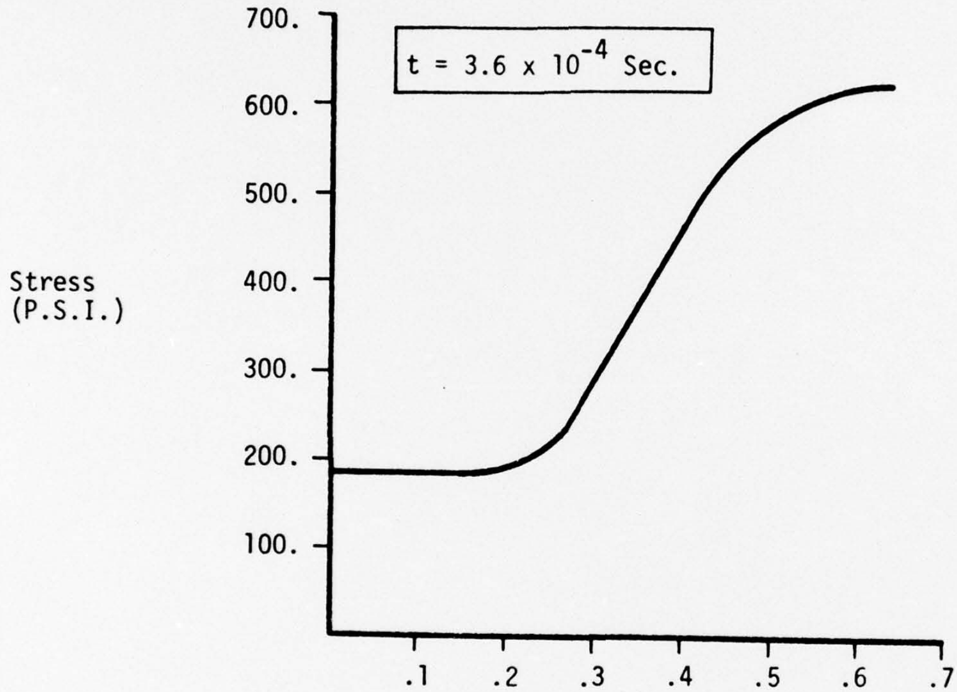


Figure 5.20 Reflection of a Nonlinear Wave by the Parabolic Regularization Method (α = 0.8).



(c)

Distance From the Free End of the Bar (In.)

Figure 5.20 (cont.) Reflection of a Nonlinear Wave by the Parabolic Regularization Method (α = 0.8).

CHAPTER VI

A THEORY OF PARABOLIC REGULARIZATION/SHOCK SMEARING METHODS FOR WAVE AND SHOCK PROPAGATION

VI.1 Introduction. In this chapter we take another point of view in solving problems of shock propagation. From the results of Chapter IV it is clear that the nonconvergence of the standard Galerkin method for shocks is caused by the lack of regularity of the exact solution. The problem is one of approximation, not of stability. In this chapter we introduce an alternate method for the calculation of shock waves. This method produces convergence to shock wave solutions by using an approximation which has greatly increased smoothness over the exact solution but still converges to this exact solution. We construct this approximation by appending to the original functional certain regularizing terms which make the approximation parabolic rather than hyperbolic. We make these regularizing terms depend on the discretization parameters, so that they go to zero in the limit as the mesh is refined.

This parabolic regularization technique is the generalization of a physically derived Lax-Wendroff type scheme (see Fost [42], Fost, Oden, and Wellford [43], and Oden [45]). It is thus in many ways similar to the shock smearing schemes used in finite difference calculations (see Richtmyer and Morton [11]). However, the contribution of the

Galerkin/parabolic regularization method is still substantial. This contribution is in both the theoretical and computational areas.

Shock smearing schemes are perpetually plagued by the problem of too much dissipation and a lack of sufficient accuracy. We use here the theoretical power of the variational method to determine the variation of the accuracy, stability, and thus the dissipation with the regularizing parameter through its dependence on the discretization parameters. Our primary tool here is a series of regularity results for an auxiliary parabolic problem to the original hyperbolic problem. Essentially, we use the regularity theory of weak partial differential equations to determine the rate at which we lose derivatives in the auxiliary problem as the regularizing parameter is decreased. Then, using these results, the regularizing parameter is optimized.

It is our conclusion that convergence can be obtained for only a limited range of variation of the regularizing parameter. Essentially the regularizing parameter cannot be made too large. For regularizing parameters in this range, we obtain a precise expression for the accuracy in terms of the regularizing parameter. We show that the regularizing parameter must be made as small as possible to minimize the approximation error. Then we determine a criteria for numerical stability of the method which determines the minimum permissible value of the regularizing parameter.

VI.2 Some Preliminiaries. In this chapter we generalize the scope of our problem. In particular, we consider Ω to be a bounded open domain in n-dimensional Euclidean space, $\mathbb{R}^n$ with boundary $\partial\Omega$.

We denote by $H^m(\Omega)$ the Sobolev space of order M. This is the space of functions whose generalized derivatives of order $\leq$ m are in $L_2(\Omega)$. The norm on $H^m(\Omega)$ is

$$\|u\|^2_{H^m(\Omega)} = \sum_{|\underset{\sim}{\alpha}|\leq m} \|D^{\underset{\sim}{\alpha}}u\|^2_{L_2(\Omega)} \tag{6.2.1}$$

Here we have used the multi-index notation; $\underset{\sim}{\alpha}$ is an ordered n-tuple of nonnegative integers: $\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, α_i = integer ≥ 0. Also, we employ the conventions,

$$X^{\underset{\sim}{\alpha}} = X_1^{\alpha_1} X_2^{\alpha_2} \ldots X_n^{\alpha_n} \; ; \; |\underset{\sim}{\alpha}| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$$

$$D^{\underset{\sim}{\alpha}}u = \left(\frac{\partial^{\alpha_1}}{\partial X_1^{\alpha_1}}\right)\left(\frac{\partial^{\alpha_2}}{\partial X_2^{\alpha_2}}\right) \ldots \frac{\partial^{\alpha_n}u}{\partial X_n^{\alpha_n}} = \frac{\partial^{|\underset{\sim}{\alpha}|}u}{\partial X_1^{\alpha_1}\partial X_2^{\alpha_2}\ldots\partial X_n^{\alpha_n}}$$

where $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n) \in \Omega$. $H^m(\Omega)$ is the completion of $C^m(\Omega)$ the space of functions with continuous derivatives of order m relative to the norm (6.2.1).

In addition, $H_o^m(\Omega)$ shall denote the completion of $C_o^\infty(\Omega)$, the space of infinitely differentiable functions with compact support in Ω, with respect to the norm (6.2.1). For a sufficiently smooth boundary $\partial\Omega$, $H_o^m(\Omega)$ is the set of functions in $H^m(\Omega)$ which satisfy homogeneous boundary conditions of order m-1; i.e., $D^{\underset{\sim}{\alpha}}u(\underset{\sim}{X}) = 0$, $\underset{\sim}{X} \in \partial\Omega$ $|\underset{\sim}{\alpha}| \leq m-1$. Note that if $u \in H_o^m(\Omega)$, u satisfies a Friedrich's inequality of the form (for some $B \geq 0$)

$$\|u\|^2_{L_2(\Omega)} \leq B \sum_{i=1}^{n} \int_\Omega \left(\frac{\partial u}{\partial X_i}\right)^2 dx \tag{6.2.2}$$

Let $u(\underset{\sim}{X},t)$ be a function defined on $\Omega \times [0,T]$ such that $u: [0,T] \to H^m(\Omega)$. We shall say that $u \in L_2(H^m(\Omega))$ if $u \in L_2(0,T)$ in the temporal variable t and $u \in H^m(\Omega)$ in the spatial variable $\underset{\sim}{X}$. The norm on $L_2(H^m(\Omega))$ is

$$\|u\|^2_{L_2(H^m(\Omega))} = \int_0^T \|u(t)\|^2_{H^m(\Omega)} \, dt \qquad (6.2.3)$$

Normally this space is denoted $L_2(a,b;H^m(\Omega))$. where $(a,b) \in \mathbb{R} = (-\infty,\infty)$, but since we shall always be dealing with the fixed time interval $(a,b) = (0,T)$, the notation $L_2(H^m(\Omega)) \equiv L_2(0,T;H^m(\Omega))$ is used. We say that $u \in L_\infty(H^m(\Omega))$ if $u \in L_\infty(0,T)$ for each $\underset{\sim}{X} \in \Omega$ and $u \in H^m(\Omega)$ for each $t \in (0,T)$. The norm on $L_\infty(H^m(\Omega))$ is

$$\|u\|_{L_\infty(H^m(\Omega))} = \sup_{0 \le t \le T} \|u(t)\|_{H^m(\Omega)} \qquad (6.2.4)$$

In addition, we shall use the symbol $(\cdot\ ,\ \cdot)_0$ to denote the L_2 inner product:

$$(u,v)_0 = \int_\Omega u\, v\, dx \quad u,v \in L_2(\Omega) \qquad (6.2.5)$$

Then a special notation for the L_2 norm is $\|u\|_0 = (u,u)_0^{1/2}$.

<u>VI.3 The Variational Problem</u>. First consider a classical nonlinear hyperbolic problem of the second order characterized as follows: Find a function $u(\underset{\sim}{X},t)$, $(\underset{\sim}{X},t) \in \Omega \times [0,T]$, such that

$$\rho \frac{\partial^2 u}{\partial t^2}(\underset{\sim}{X},t) - \underset{\sim}{\nabla} \cdot (C^2(\underset{\sim}{X},t,u)\, \underset{\sim}{\nabla}\, u(\underset{\sim}{X},t)) = f(\underset{\sim}{X},t) \quad \text{in } \Omega \times (0,T]$$

$$u(\underset{\sim}{X},0) = g_1(\underset{\sim}{X}) \quad \text{in } \Omega$$

$$\frac{\partial u(\underset{\sim}{X},0)}{\partial t} = g_2(\underset{\sim}{X}) \quad \text{in } \Omega \tag{6.3.1}$$

$$u(\underset{\sim}{X},t) = 0 \quad \text{in } \partial\Omega \times (0,T]$$

Even under reasonable assumptions on the functions $C^2(\underset{\sim}{X},t,u)$, $f(\underset{\sim}{X},t)$, $g_1(\underset{\sim}{X})$, and $g_2(\underset{\sim}{X})$, we often can be assured that a solution exists for this problem only for certain choices of T. For the moment, assume that such a solution exists for all $t \in (0,T]$. We remark that the notation $C^2(\underset{\sim}{X},t,u)$ is used in recognition of this function as the square of the intrinsic wave speed at which "disturbances" are propagated in problem (6.3.1).

It is well known that a solution to the problem (6.3.1) is also a solution of the following weaker problem: Find

$$u(\underset{\sim}{X},t) \in L_2(H_0^1(\Omega)),\ (\underset{\sim}{X},t) \in \Omega \times [0,T], \text{ such that}$$

$$(\rho \frac{\partial^2 u}{\partial t^2}, v)_0 + a(u,u,v) = (f,v)_0 \quad \forall\ v \in H^1(\Omega)$$

$$(u(\cdot,0),v)_0 = (g_1,v)_0 \quad \forall\ v \in H^1(\Omega) \tag{6.3.2}$$

$$(\frac{\partial u}{\partial t}(\cdot,0),v)_0 = (g_2,v)_0 \quad \forall\ v \in H^1(\Omega)$$

where

$$a(u,u,v) = \int_\Omega C^2(\underset{\sim}{X},t,u)\, \underset{\sim}{\nabla} u \cdot \underset{\sim}{\nabla} v \, dx \tag{6.3.3}$$

Here we complete the problem description by characterizing the real valued function $C^2(\underset{\sim}{X},t,u)$: we assume that there exist positive constants M_1, M_2, and M_3 such that

$$\begin{aligned}
&\text{(i)} \quad C^2(\underset{\sim}{X},t,u) \geq M_1 \quad \forall\ (\underset{\sim}{X},t) \in \Omega \times [0,T] \\
&\text{(ii)} \quad C^2(\underset{\sim}{X},t,u) \leq M_2 \quad \forall\ (\underset{\sim}{X},t) \in \Omega \times [0,T] \qquad (6.3.4) \\
&\text{(iii)} \quad |C^2(\underset{\sim}{X},t,u) - C^2(X,t,\bar{u})| \leq M_3\ |u - \bar{u}| \quad \forall (\underset{\sim}{X},t) \in \Omega \times [0,T]
\end{aligned}$$

These restrictions imply positiveness, boundedness, and Lipschitz continuity respectively, of the wave speed function C^2.

A <u>coercive</u> <u>property</u> of $a(u,u,v)$ can now be established through the following lemma:

<u>Lemma 6.1</u>: Let v, $\omega \in H_0^1(\Omega)$, then there exists a positive constant μ such that

$$a(v,\omega,\omega) \geq \mu\ \|\omega\|^2_{H^1(\Omega)} \qquad (6.3.5)$$

<u>Proof</u>: We have $C^2(\underset{\sim}{X},t,v) \geq M_1$. Thus

$$a(v,\omega,\omega) \geq M_1 \int_\Omega \underset{\sim}{\nabla}\omega \cdot \underset{\sim}{\nabla}\omega\ dx$$

But by use of Sobolev's imbedding theorem (see, for example [16]) and Freidrick's inequality (6.2.2).

$$\|\omega\|^2_{H^1(\Omega)} = \|\omega\|_0^2 + \sum_{i=1}^{n} \left\| \frac{\partial\omega}{\partial X_i} \right\|_0^2$$

$$\leq B \sum_{i=1}^{n} \left\| \frac{\partial \omega}{\partial X_i} \right\|_o^2 + \sum_{i=1}^{n} \left\| \frac{\partial \omega}{\partial X_i} \right\|_o^2$$

$$= \frac{M_1}{\mu} \sum_{i=1}^{n} \left\| \frac{\partial \omega}{\partial X_i} \right\|_o^2$$

$$\leq \frac{1}{\mu} a(v,\omega,\omega)$$

where $\mu = M_1/(B + 1)$. ■

Now suppose we identify a finite dimensional subspace M of $H_o^1(\Omega)$. Then the semidiscrete Galerkin approximation U of the weak solution u of (6.3.2) is that $U \in M$ such that

$$\left(\rho \frac{\partial^2 U}{\partial t^2}, V\right)_o + a(U,U,V) = (f,V)_o \quad \forall \quad V \in M$$

$$(U(\cdot,0),V)_o = (g_1,V)_o \quad \forall \quad V \in M \tag{6.3.6}$$

$$\left(\frac{\partial U}{\partial t}(\cdot,0),V\right)_o = (g_2,V)_o \quad \forall \quad V \in M$$

Under the stated assumptions, it can be shown that U is unique. We must now describe more precisely how the subspace M can be constructed in a systematic and computationally effective way. Toward this end, we establish some basic properties of the finite element method.

<u>VI.4 Finite Element Models</u>. To develop finite-element models of our problem, the region Ω is partitioned into a finite number E of disjoint open sets Ω_e called finite elements:

$$\Omega = \bigcup_{e=1}^{E} \bar{\Omega}_e ; \; \Omega_e \cap \Omega_f = 0 \quad e \neq f \tag{6.4.1}$$

Here $\bar{\Omega}_e$ is the closure of Ω_e. Within each element a set of local basis functions $\psi_N^{\underset{\sim}{\alpha}(e)}(\underset{\sim}{X})$ having the following properties is identified

$$D^{\underset{\sim}{\beta}} \psi_N^{\underset{\sim}{\alpha}(e)} (X_f^M) = \delta^{\underset{\sim}{\beta}\underset{\sim}{\alpha}} \delta_N^M \delta_f^e$$

$$\psi_N^{\underset{\sim}{\alpha}(e)} (\underset{\sim}{X}) = 0 \quad \underset{\sim}{X} \notin \Omega_e \tag{6.4.2}$$

$$\underset{\sim}{\alpha}, \underset{\sim}{\beta} \in Z_+^n ; \; M,N = 1, 2, \ldots , N_e ;$$

$$e,f = 1,2,\ldots,E ; \; |\underset{\sim}{\alpha}| \leq k$$

Here X_f^M is a nodal point labeled M in element Ω_f, $\delta^{\underset{\sim}{\beta}\underset{\sim}{\alpha}}$, δ_N^M, δ_f^e are kronecker deltas, N_e is the number of nodes in the element Ω_e. The <u>local</u> representation of a function in terms of the basis functions $\psi_N^{\underset{\sim}{\alpha}(e)}(\underset{\sim}{X})$ is

$$u_e(\underset{\sim}{X}) = \sum_{|\underset{\sim}{\alpha}| \leq k} \sum_{N=1}^{N_e} u^N_{\underset{\sim}{\alpha}(e)} \psi_N^{\underset{\sim}{\alpha}(e)} (\underset{\sim}{X}) \tag{6.4.3}$$

$$u^N_{\underset{\sim}{\alpha}(e)} = D^{\underset{\sim}{\alpha}} u_e(\underset{\sim}{X}_e^N) \tag{6.4.4}$$

and the global representation is of the form

$$U(\underset{\sim}{X}) = \bigcup_{e=1}^{E} U(\underset{\sim}{X}_e) = \sum_{|\underset{\sim}{\alpha}| \leq k} \sum_{\Delta=1}^{G} U^{\Delta}_{\underset{\sim}{\alpha}} \chi_{\Delta}^{\underset{\sim}{\alpha}}(\underset{\sim}{X}) \tag{6.4.5}$$

Here $\chi_{\Delta}^{\underset{\sim}{\alpha}}(\underset{\sim}{X})$ are global basis functions given by

$$\overset{\alpha}{\chi}_{\Delta}^{\sim}(\underset{\sim}{X}) = \bigcup_{e=1}^{E} \sum_{N=1}^{N_e} \overset{(e)N}{\Omega}_{\Delta} \overset{\alpha(e)}{\psi}_{N}^{\sim} (\underset{\sim}{X}_e) \tag{6.4.6}$$

where $\overset{(e)}{\Omega}{}^{N}_{\Delta}$ defines a Boolean transformation of the disconnected system of elements into the connected model Ω(i.e., $\overset{(e)}{\Omega}{}^{N}_{\Delta} = 1$ if node N of Ω_e coincides with node $\underset{\sim}{X}^{\Delta}$ of Ω and $\overset{(e)}{\Omega}{}^{N}_{\Delta} = 0$ if otherwise).

Now the set of functions $\{\overset{\alpha}{\chi}_{\Delta}^{\sim}(\underset{\sim}{X})\}_{\Delta=1}^{G}$; $|\underset{\sim}{\alpha}| \leq k$ defines a finite dimensional subspace of $H^1(\Omega)$ which we denote by $S_h^k(\Omega)$. Here h is the mesh parameter of the finite element mesh. For economy in notation, we shall relabel the global basis functions $\overset{\alpha}{\chi}_{\Delta}^{\sim}(\underset{\sim}{X})$ as $\Phi_N(\underset{\sim}{X})$, N = 1, 2, ..., N_o. Then the global representation is of the form

$$U(\underset{\sim}{X}) = \sum_{N=1}^{N_o} U^N \Phi_N(\underset{\sim}{X}) \tag{6.4.7}$$

The functions $\{\Phi_N\}_{N=1}^{N_o}$ form a basis for the subspace $S_h^k(\Omega)$, and the subscript h is used to designate that $S_h^k(\Omega)$ depends upon the conventional finite-element mesh parameter h; that is, if

$$h_e = \mathrm{dia}(\bar{\Omega}_e)$$

then

$$h = \max_{1 \leq e \leq E} \{h_e\}$$

The finite-element subspaces $S_h^k(\Omega)$ will be used for the space in the Galerkin approximation (6.3.6), and it shall be assumed that $S_h^k(\Omega)$ has the following properties (Cf. [62], [63], [76]):

(i) Let $P_j(\Omega)$ be the space of polynomials of degree j on Ω. Then there exists an integer k such that $p(\underset{\sim}{X}) \in P_j(\Omega)$ is in $S_h^k(\Omega)$ as long as $j \leq k$.

(ii) Let $h \to 0$ uniformly (i.e., for each refinement of the mesh let the radius ρ_e of the largest sphere that can be inscribed in Ω_e be proportional to h_e). Then there is a constant K independent of h such that

$$\underset{W \in S_h(\Omega)}{\mathrm{INF}} \; \| u-W \|_{H^m(\Omega)} \leq K h^{k+1-m} \|u\|_{H^{k+1}(\Omega)} \tag{6.4.8}$$

(iii) $S_h(\Omega)$ satisfies an inverse hypothesis [73] of the following form: there exists a constant C* independent of h such that

$$\|V\|_{H^j(\Omega)} \leq C^* h^{-j} \|V\|_{L_2(\Omega)} \quad \forall \;\; V \in S_h(\Omega),\; j \leq k+1 \tag{6.4.9}$$

The finite-element Galerkin model is formulated by setting $M = S_h(\Omega)$ and letting $V = \Phi_N$, $N = 1, \ldots, N_o$ in (6.3.6).

$$\left.\begin{aligned} &\left(\rho \frac{\partial^2 U}{\partial t^2}, \Phi_N\right)_o + a(U,U,\Phi_N) = (f,\Phi_N)_o \\ &(U(\cdot,0),\Phi_N)_o = (g_1,\Phi_N)_o \\ &\left(\frac{\partial U}{\partial t}(\cdot,0),\Phi_N\right)_o = (g_2,\Phi_N)_o \end{aligned}\right\} \quad \begin{aligned} &\Phi_N \in S_h(\Omega) \\ &N = 1,\ldots,N_o \end{aligned} \tag{6.4.10}$$

These equations describe a system of nonlinear, second-order ordinary differential equations in the coefficients $A^N(t)$ of the Galerkin approximation,

$$U(\underset{\sim}{X},t) = \sum_{N=1}^{N_o} A^N(t)\Phi_N(\underset{\sim}{X}) \tag{6.4.11}$$

In practical calculations, we must, of course, also identify temporal approximations so as to numerically integrate the equations in time.

VI.5 The Temporal Discretization. Let P be a partition of the time domain $[0,T]$ of the form $\{t_o, t_1, \ldots, t_N\}$ where $0 \leq t_o \leq t_1 < \ldots < t_N = T$ and $t_{n+1} - t_n = \Delta t$ for $0 \leq n \leq N-1$. The values of the dependent variable $U(t)$ at the points of the partition P are denoted by $\{U^n\}_{n=0}^{N}$.

In order to construct a Lax-Wendroff type approximation, we initially expand U^n and $\dot{U}^n$ in Taylor's series expansions. Setting $Y^n = \dot{U}^n$ for clarity, we have

$$U^{n+1} = U^n + \Delta t\, Y^n + \frac{\Delta t^2}{2}\ddot{U}^n + O(\Delta t^3) \tag{6.5.1}$$

$$Y^{n+1} = Y^n + \Delta t\ddot{U}^n + \frac{\Delta t^2}{2}\dddot{U}^n + O(\Delta t^3) \tag{6.5.2}$$

where for example $\dot{X}^n \equiv \left.\frac{\partial X}{\partial t}\right|_{t=n\Delta t}$. From (6.5.2)

$$\ddot{U}^n = \frac{Y^{n+1} - Y^n}{\Delta t} - \frac{\Delta t}{2}\dddot{U}^n + O(\Delta t^3) \tag{6.5.3}$$

Differentiating (6.4.10) with respect to time and evaluating the resulting expression at $t = n\Delta t$

$$(\rho U^n, V)_o + a(U^n, Y^n, V) + (\frac{\partial}{\partial t} c^2(\underset{\sim}{X}, t, U^n)\underset{\sim}{\nabla} U^n, \underset{\sim}{\nabla} V)_o \tag{6.5.4}$$

$$= (\frac{\partial}{\partial t} f(\underset{\sim}{X}, t), V)_o \quad \forall \quad V \in S_h(\Omega)$$

Introducing (6.5.1) into (6.4.10) evaluated at $t = n\Delta t$, we get

$$\left(\rho \frac{U^{n+1} - U^n}{\Delta t}, V\right)_0 - (\rho Y^n, V)_0 + \frac{\Delta t}{2} a(U^n, U^n, V) = \frac{\Delta t}{2} (f,V)_0, \quad \forall \ V \in S_h(\Omega) \tag{6.5.5}$$

For convenience, we assume that the last two terms in (6.5.4) are negligible compared to the first two. Then introducing (6.5.3) into (6.4.10), using (6.5.4) and neglecting terms of order Δt^3 or higher, we get

$$\left(\rho \frac{Y^{n+1} - Y^n}{\Delta t}, V\right)_0 + \frac{\Delta t}{2} a(U^n, Y^n, V) + a(U^n, U^n, V) = (f,V)_0 \quad \forall \ V \in S_h(\Omega) \tag{6.5.6}$$

Equations (6.5.5) and (6.5.6) define a <u>nonlinear finite-element/Lax-Wendroff Approximation</u> for the second order hyperbolic problem (6.3.1).

The natural generalization of the nonlinear finite-element/ Lax-Wendroff scheme is the <u>parabolic regularization method</u> defined by

$$\left(\rho \frac{U^{n+1} - U^n}{\Delta t}, V\right)_0 - (\rho Y^n, V)_0 + \frac{\Delta t^\alpha}{2} a(U^n, U^n, V) = \frac{\Delta t^\alpha}{2} (f,V)_0, \ \forall \ V \in S_h(\Omega) \tag{6.5.7}$$

and

$$\left(\rho \frac{Y^{n+1} - Y^n}{\Delta t}, V\right)_0 + \frac{\Delta t^\alpha}{2} a(U^n, Y^n, V) + a(U^n, U^n, V) = (f,V)_0 \quad \forall \ V \in S_h(\Omega) \tag{6.5.8}$$

where $\alpha > 0$.

VI.6 Regularity. In this section we develop regularity results for a system of equations of the form

$$(\rho \frac{\partial\tilde{u}}{\partial t},v)_o - (\rho\tilde{y},v)_o + \varepsilon a(\tilde{u},\tilde{u},v) = 0 \quad \forall \quad v \in H^1(\Omega)$$
$$(\rho \frac{\partial\tilde{y}}{\partial t},v)_o + a(\tilde{u},\tilde{u},v) + \varepsilon a(\tilde{u},\tilde{y},v) = 0 \quad \forall \quad v \in H^1(\Omega) \tag{6.6.1}$$

where ε is a real parameter ≥ 0. This system is useful in studying the convergence and accuracy of the parabolic regularization approximation described previously. We assume that $|\partial^m C^2(\underset{\sim}{x},t,\tilde{u})/\partial t^m|$ is bounded by positive constant M_4 for all $m \in Z_+$ and $\tilde{u} \in L_2(H^1(\Omega))$ which clearly corresponds to the case described by (6.6.1) when $\varepsilon = 0$.

Initially we examine the equation $(6.6.1)_2$. Let $v = \tilde{y}$ in $(6.6.1)_2$. Then

$$(\rho \frac{\partial\tilde{y}}{\partial t},\tilde{y})_o + a(\tilde{u},\tilde{u},\tilde{y}) + \varepsilon a(\tilde{u},\tilde{y},\tilde{y}) = 0 \tag{6.6.2}$$

Estimating the terms in (6.6.2) using the Cauchy-Schwarz inequality, Lemma 6.1, the definition of the L_2 norm, and the elementary inequality $ab \leq \frac{\alpha}{2} a^2 + \frac{1}{2\alpha} b^2$ for $\alpha > 0$ (henceforth to be called inequality E), we see that

$$\frac{1}{2} \frac{d}{dt} (\|\rho^{1/2}\tilde{y}\|_o^2) + \mu\varepsilon\|\tilde{y}\|^2_{H^1(\Omega)} \leq \frac{C}{2\varepsilon\eta} \|\tilde{u}\|^2_{H^1(\Omega)}$$
$$+ \frac{C\varepsilon\eta}{2} \|\tilde{y}\|_{H^1(\Omega)} + \|\tilde{y}\|_o^2 \tag{6.6.3}$$

This leads to a regularity result of the form

$$\|\tilde{y}\|_{L_2(H^1(\Omega))} \leq C\{\frac{1}{\varepsilon}\|\tilde{u}\|_{L_2(H^1(\Omega))} + \frac{1}{\sqrt{\varepsilon}}\|\tilde{y}(0)\|_o\} \tag{6.6.4}$$

By combining the two equations in (6.6.1), we find that

$$(\rho\frac{\partial^2\tilde{u}}{\partial t^2},v)_o + a(\tilde{u},\tilde{u},v) + \varepsilon a(\tilde{u},\frac{\partial\tilde{u}}{\partial t},v) + \varepsilon a(\tilde{u},\tilde{y},v)$$

$$+ \varepsilon(\frac{\partial}{\partial t}C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\tilde{u},\underset{\sim}{\nabla}v)_o = 0 \quad \forall\ v \in H^1(\Omega) \tag{6.6.5}$$

Now let $v = \frac{\partial\tilde{u}}{\partial t}$ and ρ = const. in (6.6.5). Then

$$\frac{1}{2}\frac{d}{dt}(\rho\|\frac{\partial\tilde{u}}{\partial t}\|_o^2) + a(\tilde{u},\tilde{u},\frac{\partial\tilde{u}}{\partial t}) + \varepsilon a(\tilde{u},\frac{\partial\tilde{u}}{\partial t},\frac{\partial\tilde{u}}{\partial t}) + \varepsilon a(\tilde{u},\tilde{y},\frac{\partial\tilde{u}}{\partial t})$$

$$+ \varepsilon(\frac{\partial}{\partial t}C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\tilde{u},\underset{\sim}{\nabla}\frac{\partial\tilde{u}}{\partial t})_o = 0 \tag{6.6.6}$$

Using a procedure similar to the one used in equation (6.6.4), we find that

$$\frac{1}{2}\frac{d}{dt}(\rho\|\frac{\partial\tilde{u}}{\partial t}\|_o^2) + \varepsilon\mu\|\frac{\partial\tilde{u}}{\partial t}\|^2_{H^1(\Omega)} \leq \frac{C_1}{\varepsilon}\{\|\tilde{u}\|^2_{H^1(\Omega)}$$

$$+ \varepsilon^2\|\tilde{y}\|^2_{H^1(\Omega)} + \varepsilon^2\|\tilde{u}\|^2_{H^1(\Omega)}\}$$

$$+ C_2\varepsilon\|\frac{\partial\tilde{u}}{\partial t}\|^2_{H^1(\Omega)} + \|\frac{\partial\tilde{u}}{\partial t}\|_o^2$$

Then integrating and applying the Gronwall inequality, we find that

$$\|\frac{\partial\tilde{u}}{\partial t}\|_{L_2(H^1(\Omega))} \leq C\{\frac{1}{\varepsilon}\|\tilde{u}\|_{L_2(H^1(\Omega))} + \frac{1}{\sqrt{\varepsilon}}\|\tilde{y}(0)\|_o + \frac{1}{\sqrt{\varepsilon}}\|\frac{\partial\tilde{u}}{\partial t}(0)\|_o\} \tag{6.6.7}$$

Now differentiating $(6.6.1)_2$ with respect to time

$$(\rho\frac{\partial^2\tilde{y}}{\partial t^2},v)_0 + a(\tilde{u},\frac{\partial\tilde{u}}{\partial t},v) + (\frac{\partial}{\partial t}\,C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\tilde{u},\underset{\sim}{\nabla}v)_0$$

$$+ \varepsilon a(\tilde{u},\frac{\partial\tilde{y}}{\partial t},v) + \varepsilon(\frac{\partial}{\partial t}\,C^2(\underset{\sim}{X},t,\tilde{u}),\underset{\sim}{\nabla}\tilde{y},\underset{\sim}{\nabla}v) = 0$$

Now setting $v = \frac{\partial\tilde{y}}{\partial t}$ and using the procedure which has become standard we can show, using (6.6.4) and (6.6.7), that

$$\|\frac{\partial\tilde{y}}{\partial t}\|_{L_2(H^1(\Omega))} \leq C\{\frac{1}{\varepsilon^2}\,\|\tilde{u}\|_{L_2(H^1(\Omega))} + \frac{1}{\varepsilon^{3/2}}\,\|\frac{\partial\tilde{u}}{\partial t}(0)\|_0$$

$$+ \frac{1}{\varepsilon^{3/2}}\,\|\tilde{y}(0)\|_0 + \frac{1}{\varepsilon^{1/2}}\,\|\frac{\partial\tilde{y}}{\partial t}(0)\|_0\} \qquad (6.6.8)$$

Differentiating (6.6.5) with respect to time, we get

$$(\rho\frac{\partial^3\tilde{u}}{\partial t^3},v)_0 + a(\tilde{u},\frac{\partial\tilde{u}}{\partial t},v) + \varepsilon a(\tilde{u},\frac{\partial^2\tilde{u}}{\partial t^2},v)$$

$$+ \varepsilon(\frac{\partial}{\partial t}\,C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\,\frac{\partial\tilde{u}}{\partial t},\,\underset{\sim}{\nabla}v)_0 + \varepsilon a(\tilde{u},\frac{\partial\tilde{y}}{\partial t},v)$$

$$+ \varepsilon(\frac{\partial}{\partial t}\,C^2(\underset{\sim}{X},t,\tilde{u}),\tilde{y},v)_0 + (\frac{\partial}{\partial t}\,C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\tilde{u},\underset{\sim}{\nabla}v)_0$$

$$+ \varepsilon(\frac{\partial}{\partial t}\,C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\frac{\partial}{\partial t}\,\tilde{u},\underset{\sim}{\nabla}v)_0 + \varepsilon(\frac{\partial^2 u}{\partial t^2}\,C^2(\underset{\sim}{X},t,\tilde{u})\underset{\sim}{\nabla}\tilde{u},\underset{\sim}{\nabla}v)_0 = 0 \qquad (6.6.9)$$

Now setting $v = \frac{\partial^2\tilde{u}}{\partial t^2}$, we get, using the standard procedure

$$\|\frac{\partial^2\tilde{u}}{\partial t^2}\|_{L_2(H^1(\Omega))} \leq C\{\frac{1}{\varepsilon^2}\|\tilde{u}\|_{L_2(H^1)} + \frac{1}{\varepsilon^{3/2}}\,\|\frac{\partial\tilde{u}}{\partial t}(0)\|_0$$

$$+ \frac{1}{\varepsilon^{3/2}} \|\tilde{y}(0)\|_o + \frac{1}{\varepsilon^{1/2}} \|\frac{\partial \tilde{y}}{\partial t}(0)\|_o$$

$$+ \frac{1}{\varepsilon^{1/2}} \|\frac{\partial^2 \tilde{u}}{\partial t^2}(0)\|_o\} \tag{6.6.10}$$

Expression for the higher derivatives can be obtained by using similar methods. We state the general result in terms of a theorem:

Theorem 6.1. Let $|\partial^m C^2(\underset{\sim}{X},t,\tilde{u})/\partial t^m| \leq M_4$ for $0 \leq m \leq i+1$. Then, if $\partial^\ell \tilde{u}/\partial t^\ell(0) \in L_2(\Omega)$ for $0 \leq \ell \leq i+1$ and $\partial^\ell \tilde{y}/\partial t^\ell(0) \in L_2(\Omega)$ for $0 \leq \ell \leq i$, the regularity of the solution $(\tilde{u},\tilde{y})$ to (6.6.1) is governed by (for some positive constant C)

$$\|\frac{\partial^i \tilde{y}}{\partial t^i}\|_{L_2(H^1(\Omega))} \leq C\{\frac{1}{\varepsilon^{i+1}} \|\tilde{u}\|_{L_2(H^1(\Omega))}$$

$$+ \sum_{k=0}^{i-1} \frac{1}{\varepsilon^{i-k+1/2}} [\|\frac{\partial^k \tilde{y}}{\partial t^k}(0)\|_o + \|\frac{\partial^{k+1} \tilde{u}}{\partial t^{k+1}}(0)\|_o]$$

$$\frac{1}{\varepsilon^{1/2}} \|\frac{\partial^i \tilde{y}}{\partial t^i}(0)\|_o\} \tag{6.6.11}$$

$$\|\frac{\partial^{i+1} \tilde{u}}{\partial t^{i+1}}\|_{L_2(H^1(\Omega))} \leq C\{\frac{1}{\varepsilon^{i+1}} \|\tilde{u}\|_{L_2(H^1(\Omega))}$$

$$+ \sum_{k=0}^{i} \frac{1}{\varepsilon^{i-k+1/2}} [\|\frac{\partial^k \tilde{y}}{\partial t^k}(0)\|_o + \|\frac{\partial^{k+1} \tilde{u}}{\partial t^{k+1}}(0)\|_o\} \tag{6.6.12}$$

■

If we need estimates for these temporal derivatives in other Sobolev norms, we use a similar procedure. Initially we assume that $(\tilde{u},\tilde{y})$ the solution to (6.6.1), is periodic on the boundary (this assumption is implicit in our formulation henceforth). Then it is possible to show that

$$\|\tilde{y}\|_{L_2(H^\ell(\Omega))} \le C\{\frac{1}{\varepsilon}\|\tilde{u}\|_{L_2(H^\ell(\Omega))} + \frac{1}{\sqrt{\varepsilon}}\|\tilde{y}(0)\|_{H^{\ell-1}(\Omega)}\} \quad \ell \ge 2 \qquad (6.6.13)$$

and that

$$\|\tilde{u}\|_{L_2(H^n(\Omega))} \le C \quad \frac{1}{\varepsilon^{\frac{3(n-1)}{2}}}\|\tilde{u}\|_{L_2(H^1(\Omega))}$$

$$+ \sum_{k=0}^{n-2} \frac{1}{\varepsilon^{\frac{2+3k}{2}}} \|\tilde{y}(0)\|_{H^{n-k-2}(\Omega)}$$

$$+ \sum_{k=0}^{n-2} \frac{1}{\varepsilon^{\frac{1+3k}{2}}} \|\tilde{u}(0)\|_{H^{n-k-1}(\Omega)}\} \quad n \ge 2 \qquad (6.6.14)$$

Then it can be shown that (6.6.13) and (6.6.14) lead to the more general estimate.

<u>Theorem 6.2</u>. Let $|\partial^m C^2(\underset{\sim}{x},t,\tilde{u})/\partial t^m| \le M_4$, $|\partial^j C^2(\underset{\sim}{x},t,\tilde{u})/\partial x_k^j| \le M_5$ for $1 \le k \le n$; $0 \le m \le i+1$; and $0 \le j \le n-1$. Then if $\partial^\ell\tilde{u}/\partial t^\ell$ $H^{n-1}(\Omega)$ for $0 \le \ell \le i+1$ and $\partial^r\tilde{y}/\partial t^r(0) \in H^{n-1}(\Omega)$ for $0 \le r \le i$, the regularity of the solution $(\tilde{u},\tilde{y})$ to (6.6.1) is given for some positive constant C by

$$\left\|\frac{\partial^i\tilde{y}}{\partial t^i}\right\|_{L_2(H^n(\Omega))} \le C\left\{\frac{1}{\varepsilon^{\frac{2i+3n-1}{2}}}\|\tilde{u}\|_{L_2(H^1(\Omega))}\right.$$

$$+\sum_{k=0}^{n-2}\frac{1}{\varepsilon^{\frac{2i+4+3k}{2}}}\|\tilde{y}(0)\|_{H^{n-k-2}(\Omega)}$$

$$+\sum_{k=0}^{n-2}\frac{1}{\varepsilon^{\frac{2i+3+3k}{2}}}\|\tilde{u}(0)\|_{H^{n-k-1}(\Omega)}$$

$$+\sum_{k=0}^{i-1}\frac{1}{\varepsilon^{i-k+1/2}}\left[\left\|\frac{\partial^k\tilde{y}}{\partial t^k}(0)\right\|_{H^{n-1}(\Omega)}+\left\|\frac{\partial^{k+1}\tilde{u}}{\partial t^{k+1}}(0)\right\|_{H^{n-1}(\Omega)}\right]$$

$$\left.+\frac{1}{\varepsilon^{1/2}}\left\|\frac{\partial^i\tilde{y}}{\partial t^i}(0)\right\|_{H^{n-1}(\Omega)}\right\} \quad \begin{matrix} i\ge 0 \\ n\ge 2\end{matrix} \tag{6.6.15}$$

$$\left\|\frac{\partial^{i+1}\tilde{u}}{\partial t^{i+1}}\right\|_{L_2(H^n(\Omega))} \le C\left\{\frac{1}{\varepsilon^{\frac{2i+3n-1}{2}}}\|\tilde{u}\|_{L_2(H^1(\Omega))}\right.$$

$$+\sum_{k=0}^{n-2}\frac{1}{\varepsilon^{\frac{2i+4+3k}{2}}}\|\tilde{y}(0)\|_{H^{n-k-2}(\Omega)}$$

$$+\sum_{k=0}^{n-2}\frac{1}{\varepsilon^{\frac{2i+3+3k}{2}}}\|\tilde{u}(0)\|_{H^{n-k-1}(\Omega)}$$

$$+\sum_{k=0}^{i}\frac{1}{\varepsilon^{i-k+1/2}}\left[\left\|\frac{\partial^k\tilde{y}}{\partial t^k}(0)\right\|_{H^{n-1}(\Omega)}\right.$$

$$+ \left\| \frac{\partial^{k+1}\tilde{u}}{\partial t^{k+1}}(0) \right\|_{H^{n-1}(\Omega)}]\} \quad \begin{array}{l} i \geq 0 \\ n \geq 2 \end{array} \tag{6.6.16}$$

■

Estimates in terms of the L_∞ norm can also be obtained. For instance

Theorem 6.3. Let $|\partial^m C^2(\underset{\sim}{x},t,\tilde{u})/\partial t^m| \leq M_4$, $|\partial^j C^2(\underset{\sim}{x},t,\tilde{u})/\partial x_k^j| \leq M_5$ for $1 \leq k \leq n$; $0 \leq m \leq i+1$; and $0 \leq j \leq n-1$. Then if $\partial^\ell \tilde{u}/\partial t^\ell(0) \in H^{n-1}(\Omega)$ for $0 \leq \ell \leq i+1$ and $\partial^r \tilde{y}/\partial t^r(0) \in H^{n-1}(\Omega)$ for $0 \leq r \leq i$, the regularity of the solution $(\tilde{u},\tilde{y})$ to (6.6.1) for some positive constant C is given by

$$\left\| \frac{\partial^i \tilde{y}}{\partial t^i} \right\|_{L_\infty(H^{n-1}(\Omega))} \leq C\{ \frac{1}{\varepsilon^{\frac{2i+3n-2}{2}}} \|\tilde{u}\|_{L_2(H^1(\Omega))}$$

$$+ \sum_{k=0}^{n-2} \frac{1}{\varepsilon^{\frac{2i+3+3k}{2}}} \|\tilde{y}(0)\|_{H^{n-k-2}(\Omega)}$$

$$+ \sum_{k=0}^{n-2} \frac{1}{\varepsilon^{\frac{2i+2+3k}{2}}} \|\tilde{u}(0)\|_{H^{n-k-1}(\Omega)}$$

$$+ \sum_{k=0}^{i-1} \frac{1}{\varepsilon^{i-k}} [\left\| \frac{\partial^k \tilde{y}}{\partial t^k}(0) \right\|_{H^{n-1}(\Omega)}$$

$$+ \left\| \frac{\partial^{k+1} u}{\partial t^{k+1}}(0) \right\|_{H^{n-1}(\Omega)}]$$

$$+ \left\| \frac{\partial^i \tilde{y}}{\partial t^i}(0) \right\|_{H^{n-1}(\Omega)} \} \quad \begin{array}{l} i \geq 0 \\ n \geq 2 \end{array}$$

$$\left\| \frac{\partial^{i+1}\tilde{u}}{\partial t^{i+1}} \right\|_{L_\infty(H^{n-1}(\Omega))} \le C\Big\{\frac{1}{\varepsilon^{\frac{2i+3n-2}{2}}} \|\tilde{u}\|_{L_2(H^1(\Omega))}$$

$$+ \sum_{k=0}^{n-2} \frac{1}{\varepsilon^{\frac{2i+3+3k}{2}}} \|\tilde{y}(0)\|_{H^{n-k-2}(\Omega)}$$

$$+ \sum_{k=0}^{n-2} \frac{1}{\varepsilon^{\frac{2i+2+3k}{2}}} \|\tilde{u}(0)\|_{H^{n-k-1}(\Omega)}$$

$$+ \sum_{k=0}^{i} \frac{1}{\varepsilon^{i-k}} \Big[\left\| \frac{\partial^{k}\tilde{y}}{\partial t^{k}}(0) \right\|_{H^{n-1}(\Omega)}$$

$$+ \left\| \frac{\partial^{k+1}\tilde{u}}{\partial t^{k+1}}(0) \right\|_{H^{n-1}(\Omega)} \Big]\Big\} \quad \begin{matrix} i \ge 0 \\ n \ge 2 \end{matrix} \qquad \blacksquare$$

<u>VI.7 Approximation Theory Results and The Gronwall Lemma</u>. Certain approximation theory results are reviewed in this section to provide a complete theory of convergence. These results will be presented as a series of known lemmas, the proofs of which can be found in the literature cited.

Suppose we define an element w of the subspace $S_h(\Omega)$ through the weighted $H^1(\Omega)$ projection introduced by Wheeler [48]. Then W satisfied

$$a(u,u - W,V) = 0 \quad \forall \quad V \in S_h(\Omega) \qquad (6.7.1)$$

Let E denote the spatial projection error,

$$E = u - W \tag{6.7.2}$$

Then the behavior of E and its time derivatives in various norms is given in the following lemma:

Lemma 6.2. Let $u, \frac{\partial u}{\partial t} \in L_\infty(H^{k+1}(\Omega))$ and $\frac{\partial^2 u}{\partial t^2} \in L_2(H^{k+1}(\Omega))$. Then there exists a constant C, independent of the discretization parameters, such that

$$\|E\|_{L_\infty(L_2(\Omega))} + \left\|\frac{\partial E}{\partial t}\right\|_{L_\infty(L_2(\Omega))} + \left\|\frac{\partial^2 E}{\partial t^2}\right\|_{L_2(L_2(\Omega))}$$

$$\leq C\{h^{k+1}\|u\|_{L_\infty(H^{k+1}(\Omega))} + h^{k+1}\left\|\frac{\partial u}{\partial t}\right\|_{L_\infty(H^{k+1}(\Omega))}$$

$$+ h^{k+1}\|u\|_{L_2(H^{k+1}(\Omega))}\} \tag{6.7.3}$$

■

This Lemma has been established, for example, by Wheeler [48].

The second Lemma is the discrete version of the classical Gronwell inequality (Cf. Lees [72]).

Lemma 6.3. (The Discrete Gronwall Inequality) If $\phi(t)$ and $\psi(t)$ are nonnegative functions with $\psi(t)$ nondecreasing, and

$$\phi(N\Delta t) \leq \psi(N\Delta t) + C\Delta t \sum_{n=0}^{N-1} \phi(n\Delta t) \tag{6.7.4}$$

Then

$$\phi(N\Delta t) \leq \psi(N\Delta t)e^{CN\Delta t} \tag{6.7.5}$$

■

We can now pass on to the investigation of the linear parabolic regularization scheme.

VI.8 The Linear Parabolic Regularization Approximation. In this section of the paper we consider the approximation of the linearized version of (6.3.2) by the corresponding linearized versions of (6.5.7) and (6.5.8). To simplify the calcuations, it is assumed that $f = 0$. However, the method presented here is in no way restricted to this case.

It is possible to split up the second order equation (6.3.2) into two coupled first order equations. This is carried out by defining the new variable $y = \frac{\partial u}{\partial t}$. Then (6.3.2) is fully equivalent to the system

$$\begin{aligned} &(\rho\frac{\partial y}{\partial t},v)_0 + a(u,v) = 0 \quad \forall \quad v \in H^1(\Omega) \\ &(\rho\frac{\partial u}{\partial t},v)_0 - (\rho y,v)_0 = 0 \quad \forall \quad v \in H^1(\Omega) \end{aligned} \tag{6.8.1}$$

Thus, when we discuss the convergence of the parabolic regularization scheme, we mean convergence to the solution of a problem (6.8.1) which is equivalent to (6.3.2).

Now we pose an auxiliary problem. Let $(\tilde{u},\tilde{y})$ be the solution to the system

$$\begin{aligned} &(\rho\frac{\partial \tilde{y}}{\partial t},v)_0 + a(\tilde{u},v) + \frac{\Delta t^{*\alpha}}{2}\, a(\tilde{y},v) = 0 \quad \forall \quad v \in H^1(\Omega) \\ &(\rho\frac{\partial \tilde{u}}{\partial t},v)_0 - (\rho\tilde{y},v) + \frac{\Delta t^{*\alpha}}{2}\, a(\tilde{u},v) = 0 \quad \forall \quad v \in H^1(\Omega) \end{aligned} \tag{6.8.2}$$

We obtain an approximate auxiliary problem by introducing the forward difference operator in (6.8.2). Then the solution to the approximate auxiliary problem $(\tilde{U}^n, \tilde{Y}^n)$ satisfies

$$\left(\rho\frac{\tilde{Y}^{n+1} - \tilde{Y}^n}{\Delta t}, V\right) + a(\tilde{U}^n, V)$$

$$+ \frac{\Delta t^{*\alpha}}{2} a(\tilde{Y}^n, V) = 0 \quad \forall \quad V \in S_h(\Omega)$$

and

$$\left(\rho\frac{\tilde{U}^{n+1} - \tilde{U}^n}{\Delta t}, V\right)_o - (\rho\tilde{Y}^n, V)_o + \frac{\Delta t^{*\alpha}}{2} a(\tilde{U}^n, V) = 0 \quad \forall \quad V \in S_h(\Omega) \qquad (6.8.3)$$

To demonstrate the convergence and determine the rate of convergence, we will show that, for stable schemes,

$$U^n \xrightarrow[\Delta t \to \Delta t^*]{} \tilde{U}^n \xrightarrow[\substack{\Delta t \to 0 \\ h \to 0}]{} \tilde{u} \xrightarrow[\Delta t^* \to 0]{} u$$

$$Y^n \xrightarrow[\Delta t \to \Delta t^*]{} \tilde{Y}^n \xrightarrow[\substack{\Delta t \to 0 \\ h \to 0}]{} \tilde{y} \xrightarrow[\Delta t^* \to 0]{} y$$

and this implies that

$$U^n \xrightarrow[\substack{\Delta t \to 0 \\ h \to 0}]{} u$$

$$Y^n \xrightarrow[\substack{\Delta t \to 0 \\ h \to 0}]{} y$$

Initially, we will investigate the convergence of (6.8.3) to (6.3.2). Then using the regularity results of section VI.6, the convergence of

(6.8.2) to (6.8.1) will be determined. As a final step Δt^* in (6.8.3) will be made to approach Δt, and the convergence of (6.5.7) and (6.5.8) to (6.8.1) will be the results.

If $(6.8.2)_2$ is evaluated at time $t = n\Delta t$ and v is equated to V, then it can be seen that

$$(\rho\frac{\partial\tilde{u}_n}{\partial t},V)_o - (\rho\tilde{y}_n,V)_o + \frac{\Delta t^{*\alpha}}{2} \; a(\tilde{u}_n,V) = 0, \quad \forall \quad V \in S_h(\Omega) \tag{6.8.4}$$

Now adding $(\frac{\tilde{u}_{n+1} - \tilde{u}_n}{\Delta t},V)_o$ to each side of (6.8.4) gives

$$(\rho\frac{\tilde{u}_{n+1} - \tilde{u}_n}{\Delta t},V)_o - (\rho\tilde{y}_n,V)_o + \frac{\Delta t^{*\alpha}}{2} a(\tilde{u}_n,V) = (\rho\psi_N,V)_o \quad V \in S_h(\Omega) \tag{6.8.5}$$

where

$$\psi_n = \frac{\tilde{u}_{n+1} - \tilde{u}_n}{\Delta t} - \left.\frac{\partial\tilde{u}}{\partial t}\right|_{t=n\Delta t}$$

It can be shown that

$$\psi_n = \int_{t_n}^{t_{n+1}} (t_{n+1} - t)\,\frac{\partial^3\tilde{u}}{\partial t^3}\,dt \tag{6.8.6}$$

and an index of the accumulated temporal approximation error is

$$\psi = \sum_{n=0}^{N-1} \|\psi_n\|^2_{L_2(\Omega)} = \sum_{n=0}^{N-1} \left(\int_{t_n}^{t_{n+1}} (t_{n+1} - t)^2 \left\|\frac{\partial^3\tilde{u}(t)}{\partial t^3}\right\|^2_{L_2(\Omega)} dt\right) \tag{6.8.7}$$

Using the Cauchy inequality, we get

$$\psi \leq \sum_{n=0}^{N-1} \left(\int_{t_n}^{t_{n+1}} (t_{n+1} - t)^2 dt \right) \left(\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^3 \tilde{u}(t)}{\partial t^3} \right\|_{L_2(\Omega)} dt \right)$$

$$= \frac{\Delta t^3}{3} \left\| \frac{\partial^3 \tilde{u}}{\partial t^3} \right\|^2_{L_2(L_2(\Omega))}$$

Thus

$$\Delta t \psi = \Delta t \sum_{n=0}^{N-1} \|\psi_n\|^2_{L_2(\Omega)} \leq \frac{\Delta t^4}{3} \left\| \frac{\partial^3 \tilde{u}}{\partial t^3} \right\|^2_{L_2(L_2(\Omega))} \tag{6.8.8}$$

Now set $e_n = \tilde{u} - \tilde{U}^n$ and $f_n = \tilde{y} - \tilde{Y}^n$. Then subtracting $(6.8.3)_2$ from (6.8.5) gives

$$\left(\rho \frac{e_{n+1} - e_n}{\Delta t}, V\right)_0 - (\rho f_n, V) + \frac{\Delta t^{*\alpha}}{2} a(e_n, V) = (\rho \psi_n, V)_0 \quad \forall \ V \in S_h(\Omega) \tag{6.8.9}$$

We identify elements W^n, $P^n \in S_h(\Omega)$ through the weighted $H^1(\Omega)$ projection, as was done previously,

$$\begin{aligned} a(\tilde{u}_n - W^n, V) &= 0, \quad \forall \ V \in S_h(\Omega) \\ a(\tilde{y}_n - P^n, V) &= 0, \quad \forall \ V \in S_h(\Omega) \end{aligned} \tag{6.8.10}$$

Then, we perform the normal decomposition of the approximation error e_n and f_n. Let $e_n = E_n + \mathcal{E}_n$ where $E_n = \tilde{u}_n - W^n$ and $\mathcal{E}_n = W^n - \tilde{U}^n$ and $f_n = F_n + \mathcal{F}_n$ where $F_n = \tilde{y}_n - P^n$ and $\mathcal{F}_n = P^n - \tilde{Y}^n$.

The following theorem describes the behavior of $\mathcal{E}_n$:

Theorem 6.4. Let $\frac{\partial^3 \tilde{u}}{\partial t^3} \in L_2(L_2(\Omega))$, and

$$\frac{\Delta t^{*\alpha}}{h^2} < \frac{8\rho}{C'C^{*2}}$$

where β is a positive constant. Then there exists a constant C_1 such that

$$\|E\|_{\hat{L}_\infty(L_2(\Omega))} \leq C_1\{\|E_0\|_0 + \|F\|_{\hat{L}_\infty(L_2(\Omega))} \left\|\frac{\partial E}{\partial t}\right\|_{L_2(L_2(\Omega))}$$

$$+ \|F\|_{L_\infty(L_2(\Omega))} + \Delta t^2 \left\|\frac{\partial^3 \tilde{u}}{\partial t^3}\right\|_{L_2(L_2(\Omega))}\} \qquad (6.8.11)$$

where $\|E\|_{\hat{L}_\infty(L_2(\Omega))} = \sup_{0 \leq i \leq N} \|E_i\|_0$.

Proof: Decomposing the error in (.8.9) we get

$$(\rho\delta_t E_{n+1/2}, V)_0 + \frac{\Delta t^{*\alpha}}{2} a(E_n, V) = -(\rho\delta_t E_{n+1/2}, V) + (\rho F_n, V)_0$$

$$+ (\rho F_n, V)_0 - \frac{\Delta t^{*\alpha}}{2} a(E_n, V)$$

$$+ (\rho\psi_n, V)_0 \quad \forall \quad V \in S_h(\Omega) \qquad (6.8.12)$$

We can now set $V = E_{n+1/2}$ since $E_{n+1/2} \in S_h(\Omega)$. Then, using (6.8.10), the Cauchy-Schwarz inequality, inequality E, the inverse property (6.4.9), and the coercive property (6.3.5), we get for a positive constant ψ

$$\frac{1}{\Delta t}\left[\left(1 - \frac{\Delta t^{*\alpha} C' C^{*2}}{8\rho h^2}\right) \| \rho^{1/2} E_{n+1} \|_o^2 - \left(1 + \frac{\Delta t^{*\alpha} C' C^{*2}}{8\rho h^2}\right) \| \rho^{1/2} E_n \|_o^2\right]$$

$$+ \Delta t^{*\alpha} \psi \, \| E_n \|^2_{H^1(\Omega)}$$

$$\leq - (\rho \delta_t E_{n+1/2}, E_{n+1/2})_o + (\rho F_n, E_{n+1/2})_o + (\rho F_n, E_{n+1/2})_o + (\rho \psi_n, E_{n+1/2})_o \tag{6.8.13}$$

As a <u>condition</u> <u>of</u> <u>stability</u> we require that

$$\frac{\Delta t^{*\alpha}}{h^2} \leq \frac{8\rho}{C' C^{*2}} = G$$

This implies that $g(\Delta t) = \left(1 - \frac{\Delta t^{*\alpha}}{h^2 G}\right)\left(1 + \frac{\Delta t^{*\alpha}}{h^2 G}\right)^{-1}$ is bounded above and below by positive constants. Now estimating the terms on the right-hand side of (6.8.13) using the Cauchy-Schwarz inequality and inequality E, multiplying by $\Delta t\left[1 + \frac{\Delta t^{*\alpha}}{h^2 G}\right]^{-1} g(\Delta t)^n$ (see [47]) and summing from 1 to N-1, we get for positive constants B_1, B_2, B_3

$$\| \rho^{1/2} E_N \|_o - B_1 \| \rho^{1/2} E_o \|_o^2 + B_2 \Delta t \Delta t^{*\alpha} \psi \sum_{n=0}^{N-1} \| E_n \|^2_{H^1(\Omega)}$$

$$\leq B_3\Big(\Delta t \sum_{n=0}^{N-1} \Big\{\frac{\gamma}{2} \| \rho^{1/2} \delta_t E_{n+1/2} \|_o^2 + \frac{\eta}{2} \| \rho^{1/2} F_n \|_o^2 + \frac{\xi}{2} \| \rho^{1/2} F_n \|_o^2$$

$$+ \frac{\omega}{2} \| \rho^{1/2} \psi_n \|_o^2 \Big\} + \Delta t \sum_{n=0}^{N-1} \nu \{ \| \rho^{1/2} E_{n+1} \|_o^2 + \| \rho^{1/2} E_n \|_o^2 \}\Big) \tag{6.8.14}$$

where γ, η, ξ, ω, and ν are positive constants and $\nu = \frac{1}{4\gamma} + \frac{1}{4\eta} + \frac{1}{4\xi} + \frac{1}{4\omega}$.

A useful integral representation for terms of the form $\delta_t E_{n+1/2}$ was obtained by Dupont [49]. It can be shown that since

$$\delta_t E_{n+1/2} = \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \frac{\partial E}{\partial t}\, dt$$

then

$$\Delta t \sum_{n=0}^{N-1} \| \rho^{1/2} \delta_t E_{n+1/2} \|_0^2 \le \| \rho^{1/2} \frac{\partial E}{\partial t} \|^2_{L_2(L_2(\Omega))} \qquad (6.8.15)$$

In addition

$$\sup_{1 \le i \le N} \Delta t \sum_{n=0}^{i-1} K_1 \|F_n\|_0^2 \le K_1 N \Delta t \sup_{0 \le n \le N} \|F_n\|_0^2$$

$$= K_2 \|F_n\|^2_{\hat{L}_\infty(L_2(\Omega))} \qquad (6.8.16)$$

Using the discrete Gronwall inequality (Lemma 6.3) in (6.8.14) taking the supremum over all n in the resulting expression, using (6.8.15) and (6.8.16), and introducing the temporal error term (6.8.8) we obtain the result (6.8.11). ∎

Now we examine the approximation error induced in the approximation of $(6.8.2)_1$ by $(6.8.3)_1$. If $(6.8.2)_1$ is evaluated at $t = n\Delta t$ and v is set equal to $V \in S_h(\Omega)$, then

$$(\rho \frac{\partial \tilde{y}_n}{\partial t}, V)_o + a(\tilde{u}_n, V) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{y}_n, V) = 0 \quad \forall \; V \in S_h(\Omega) \tag{6.8.17}$$

Now adding $(\tilde{y}_{n+1} - \tilde{y}_n/\Delta t, V)_o$ to each side of (6.8.17), we get

$$(\frac{\tilde{y}_{n+1} - \tilde{y}_n}{\Delta t}, V)_o + a(\tilde{u}_n, V) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{y}_n, V) = (\rho\beta_n, V)_o \; \forall \; V \in S_h(\Omega) \tag{6.8.18}$$

where

$$\beta_n = \frac{\tilde{y}_{n+1} - \tilde{y}_n}{\Delta t} - \frac{\partial \tilde{y}_n}{\partial t}$$

An estimate for the temporal error component (see the derivative of (6.8.8)) is

$$\Delta t \sum_{n=0}^{N-1} \| \beta_n \|^2_{L_2(\Omega)} \leq \frac{\Delta t^4}{3} \| \frac{\partial^3 \tilde{y}}{\partial t^3} \|^2_{L_2(L_2(\Omega))} \tag{6.8.19}$$

Now subtracting $(6.8.3)_1$ from (6.8.18), we find that for $\forall \; V$ in $S_h(\Omega)$,

$$(\rho \frac{f_{n+1} - f_n}{\Delta t}, V)_o + a(e_n, V)_o + \frac{\Delta t^{*\alpha}}{2} a(f_n, V) = (\rho\beta_n, V)_o \tag{6.8.20}$$

The approximation error f_n and e_n are then decomposed in the normal manner. The behavior of F_n is given in the following theorem:

<u>Theorem 6.5</u>. Suppose $\frac{\partial^3 \tilde{y}}{\partial t^3} \in L_2(L_2(\Omega))$ and suppose that we choose Δt and h so that

$$\frac{h^2}{\Delta t^*} < \frac{C'C^{*2}}{\psi}$$

$$\frac{\Delta t^{*2}}{\Delta th^2} < \frac{\psi^2}{2\lambda C^{*2}\zeta C'} \qquad (6.8.21)$$

$$\frac{\Delta t^{*2-2\alpha}}{\Delta th^2} < \frac{\psi^2\zeta}{2\lambda C^{*2}C'}$$

where ψ, ζ, C', C^* are positive constants and ψ,ζ are arbitrarily chosen and $\lambda = 1 - \psi h^2/C'C^{*2}\Delta t^*$. Then there exists a constant C_2 such that

$$\|F\|_{\hat{L}_\infty(L_2(\Omega))} \leq C_2\{\|F_0\|_0 + \|\frac{\partial F}{\partial t}\|_{L_2(L_2(\Omega))} + \Delta t^2 \|\frac{\partial^3\tilde{y}}{\partial t^3}\|_{L_2(L_2(\Omega))}\}$$

<u>Proof</u>: Decompose the error in (6.8.21) according to

$$(\rho\frac{F_{n+1} - F_n}{\Delta t},V)_0 + a(E_n,V)_0 + \frac{\Delta t^{*\alpha}}{2} a(F_n,V)$$

$$= - (\rho\frac{F_{n+1} - F_n}{\Delta t},V)_0 - a(E_n,V)_0 - \frac{\Delta t^{*\alpha}}{2} a(F_n,V)$$

$$+ (\rho\beta_n,V)_0 \qquad (6.8.23)$$

and take note of the following identity:

$$\delta_{t_{n+1/2}} (F,F)_0 = 2(F_n,\delta_t F_{n+1/2})_0 + \Delta t(\delta_t F_{n+1/2},\delta_t F_{n+1/2})_0 \qquad (6.8.24)$$

Setting $V = F_n$ in (6.8.23) and using (6.3.24), we get

$$\delta_{t_{n+1/2}}(\rho F,F)_o - \Delta t(\rho\delta_t F_{n+1/2},\delta_t F_{n+1/2})_o + 2a(E_n,F_n)$$

$$+ \Delta t^{*\alpha} a(F_n,F_n) = -2(\rho \frac{F_{n+1} - F_n}{\Delta t}, F_n)_o$$

$$- 2a(E_n,F_n) - \Delta t^{*\alpha} a(F_n,F_n) + 2(\rho\beta_n,F_n)_o \qquad (6.8.25)$$

Using (6.8.10), Lemma 6.1, and inequality E (for some positive constant ζ), we get

$$\delta_{t_{n+1/2}}(\rho F,F)_o - \Delta t(\rho\delta_t F_{n+1/2},\delta_t F_{n+1/2})_o - \zeta C'\|E_n\|^2_{H^1(\Omega)}$$

$$+ \Delta t^{*\alpha}\mu\|F_n\|^2_{H^1(\Omega)} = -2(\rho \frac{F_{n+1} - F_n}{\Delta t}, F_n)_o$$

$$+ \frac{C'}{\zeta}\|F_n\|^2_{H^1(\Omega)} + 2(\rho\beta_n,F_n)_o \qquad (6.8.26)$$

Now we simplify this expression by defining an auxiliary relationship. Letting $V = \delta_t F_{n+1/2}$ in (6.8.23), we get

$$\|\rho^{1/2}\delta_t F_{n+1/2}\|^2_o = -a(E_n,\delta_t F_{n+1/2}) - \frac{\Delta t^{*\alpha}}{2} a(F_n,\delta_t F_{n+1/2})$$

$$- (\rho\delta_t F_{n+1/2},\delta_t F_{n+1/2})_o + (\rho\beta_n,\delta_t F_{n+1/2})_o \qquad (6.8.27)$$

Using the Cauchy-Schwarz inequality, the inequality E, and the inverse assumption , we find that for ψ a positive constant

$$(1 - \frac{C'C^{*2}\Delta t^*}{\psi h^2})\|\rho^{1/2}\delta_t F_{n+1/2}\|^2_o$$

$$\leq \frac{C'\psi}{2\Delta t^*} \|E_n\|^2_{H^1(\Omega)} + \frac{C'\Delta t^{*2\alpha-1}}{2} \|F_n\|^2_{H^1(\Omega)}$$

$$- (\rho\delta_t F_{n+1/2}, \delta_t F_{n+1/2})_o + (\rho\beta_n, \delta_t F_{n+1/2})_o \qquad (6.8.28)$$

This implies that

$$\| \rho^{1/2}\delta_t F_{n+1/2}\|^2_o \leq \frac{C'\psi^2 h^2}{2\Delta t^*(\psi h^2 - C'C^{*2}\Delta t^*)} \|E_n\|^2_{H^1(\Omega)}$$

$$+ \frac{C'\psi^2 h^2 \Delta t^{*2\alpha-1}}{2\psi h^2 - 2C'C^{*2}\Delta t^*} \|F_n\|^2_{H^1(\Omega)} + BZ \qquad (6.8.29)$$

where

$$Z = (\rho\delta_t F_{n+1/2}, \delta_t F_{n+1/2})_o - (\rho\beta_n, \delta_t F_{n+1/2})_o$$

and

$$B = \frac{\psi h^2}{C'C^{*2}\Delta t^* - \psi h^2}$$

Using (6.8.29) to simplify (6.8.26)

$$\delta_{t_{n+1/2}} \| \rho^{1/2} F\|^2_o + [\frac{\psi^2 h^2 \Delta t}{2(1 - \frac{\psi}{C'C^{*2}} \frac{h^2}{\Delta t})C^{*2}\Delta t^{*2}} - \zeta C']\|E_n\|^2_{H^1(\Omega)}$$

$$+ [\frac{\psi^2 h^2 \Delta t}{(1 - \frac{\psi}{C'C^{*2}} \frac{h^2}{\Delta t^*})2C^{*2}\Delta t^{*2-2\alpha}}$$

$$+ \Delta t^{*\alpha}\mu - \frac{C'}{\zeta}]\|F_n\|^2_{H^1(\Omega)}$$

$$\leq -2(\rho \frac{F_{n+1} - F_n}{\Delta t}, F_n)_o + 2(\rho\beta_n, F_n)_o + \Delta t BZ \qquad (6.8.30)$$

As <u>conditions</u> <u>of</u> <u>stability</u> we require that

$$\left[\frac{\psi^2 h^2 \Delta t}{2\left(1 - \frac{\psi}{C'C^{*2}} \frac{h^2}{\Delta t}\right) C^{*2} \Delta t^{*2}} - \zeta C'\right] = \eta \geq 0 \tag{6.8.31}$$

and

$$\left[\frac{\psi^2 h^2 \Delta t}{2\left(1 - \frac{\psi}{C'C^{*2}} \frac{h^2}{\Delta t^*}\right) C^{*2} \Delta t^{*2-2\alpha}} - \frac{C'}{\zeta}\right] = \gamma \geq 0 \tag{6.8.32}$$

where, of course, ψ and ζ are the arbitrary positive constants introduced previously. Clearly (6.8.31) and (6.8.32) are satisfied if the conditions (6.8.21) are satisfied. Then estimating the term on the right-hand side of (6.8.31) using the Cauchy-Schwarz inequality and inequality E

$$\frac{1}{\Delta t}\left[\|\rho^{1/2} F_{n+1}\|_o - \|\rho^{1/2} F_n\|_o^2 + \eta \|E_n\|^2_{H^1(\Omega)} + \gamma \|F_n\|^2_{H^1(\Omega)}\right.$$

$$\leq \{(1 + \Delta t B)^2 \gamma \|\rho^{1/2} \delta_t F_{n+1/2}\|_o^2 + (1 + \Delta t B)^2 \phi \|\rho^{1/2} \beta_n\|_o^2\}$$

$$+ \xi \|\rho^{1/2} F_n\|_o^2 \tag{6.8.33}$$

where γ and ϕ are positive constants and $\xi = \frac{1}{2\gamma} + \frac{1}{2\phi}$.

Multiplying (6.8.33) by Δt, summing from 1 to N-1, applying the discrete Gronwall inequality given in Lemma 6.3, taking the supremum over n in the resulting expression, and using the temporal error term (6.8.19), we obtain the result (6.8.22). ■

We remark that the stability restrictions (6.8.21) are very severe. For a given value of the spatial discretization parameter h, $(6.8.21)_2$ and $(6.8.21)_3$ specifies that Δt cannot be too small. That is, instability can result from either the choice of too small or too large a discretization parameter.

We obtain the final error estimate for $\|E\|_{\hat{L}_\infty(L_2(\Omega))}$ by combining Theorem 6.4 and 6.5.

<u>Theorem 6.6</u>. *Suppose* $\partial^3\tilde{u}/\partial t^3$, $\partial^3\tilde{y}/\partial t^3 \in L_2(L_2(\Omega))$ and suppose that the conditions of Theorem 6.5 are satisfied. Then there exists a positive constant C_3 such that

$$\|E\|_{\hat{L}_\infty(L_2(\Omega))} \le C_3\{\|E_0\|_0 + \|F_0\|_0 + \|\frac{\partial E}{\partial t}\|_{L_2(L_2(\Omega))}$$

$$+ \|F\|_{L_\infty(L_2(\Omega))} + \|\frac{\partial F}{\partial t}\|_{L_2(L_2(\Omega))}$$

$$+ \Delta t^2\|\frac{\partial^3\tilde{u}}{\partial t^3}\|_{L_2(L_2(\Omega))} + \Delta t^2\|\frac{\partial^3\tilde{y}}{\partial t^3}\|_{L_2(L_2(\Omega))}\} \qquad (6.8.34)$$

■

Using the Theorem 6.6, Lemma 6.2, and the triangle inequality, we obtain the error estimate for e and f.

<u>Theorem 6.7</u>. Suppose that $\tilde{u},\tilde{y},\frac{\partial\tilde{u}}{\partial t} \in L_\infty(H^{k+1}(\Omega))$, $\frac{\partial\tilde{y}}{\partial t} \in L_2(H^{k+1}(\Omega))$, $\frac{\partial^3\tilde{u}}{\partial t^3} \in L_2(L_2(\Omega))$ and $\frac{\partial^3\tilde{y}}{\partial t^3} \in L_2(L_2(\Omega))$ and suppose that the conditions of Theorem 6.5 are satisfied. Then there exist positive constants C_4 and C_5 such that

$$\|e\|_{L_\infty(L_2(\Omega))} \le C_4\{\|e_o\|_o + \|f_o\|_o + h^{k+1}\|\tilde{u}\|_{L_\infty(H^{k+1}(\Omega))}$$

$$+ h^{k+1}\|\frac{\partial\tilde{u}}{\partial t}\|_{L_\infty(H^{k+1}(\Omega))} + h^{k+1}\|\tilde{y}\|_{L_\infty(H^{k+1}(\Omega))}$$

$$+ h^{k+1}\|\frac{\partial\tilde{y}}{\partial t}\|_{L_\infty(H^{k+1}(\Omega))}$$

$$+ \Delta t^2\|\frac{\partial^3\tilde{u}}{\partial t^3}\|_{L_2(L_2(\Omega))} + \Delta t^2\|\frac{\partial^3\tilde{y}}{\partial t^3}\|_{L_2(L_2(\Omega))}\} \qquad (6.8.35)$$

and

$$\|f\|_{\hat{L}_\infty(L_2(\Omega))} \le C_5\{\|f_o\|_o + h^{k+1}\|\frac{\partial\tilde{y}}{\partial t}\|_{L_2(H^{k+1}(\Omega))}$$

$$+ \Delta t^2\|\frac{\partial^3\tilde{y}}{\partial t^3}\|_{L_2(L_2(\Omega))}\} \qquad (6.8.36)$$

■

We can refine the estimate by introducing the regularity results of Theorems 6.1, 6.2. and 6.3 with ε replaced by $(\Delta t^*)^\alpha$.

Theorem 6.8. If the hypotheses of Theorem 6.7 are satisfied, then

$$\|e\|_{\hat{L}_\infty(L_2(\Omega))} \le 0\,[\frac{h^{k+1}}{(\Delta t^*)^{\alpha(\frac{3k+6}{2})}}]$$

$$+ 0\,[\frac{\Delta t^2}{(\Delta t^*)^{\frac{7\alpha}{2}}}]$$

and

$$\|f\|_{\hat{L}_\infty(L_2(\Omega))} \le O\left[\frac{h^{k+1}}{(\Delta t^*)^{\alpha(\frac{3k+4}{2})}}\right]$$

$$+ O\left[\frac{\Delta t^2}{(\Delta t^*)^{\frac{7\alpha}{2}}}\right] \quad \blacksquare$$

If we set $\Delta t^* = \Delta t$ and vary the discretization parameters so that $\Delta t/h^q = C$ where C is a positive constant, we obtain the final estimate of the rate of convergence:

Theorem 6.9. If the hypotheses of Theorem 6.7 are satisfied and $\Delta t/h^q = C$, then

$$\| u - U\|_{\hat{L}_\infty(L_2)} \le O\left[h^{\frac{2k - 3\alpha qk - 6\alpha q + 2}{2}}\right]$$

$$+ O\left[\Delta t^{\frac{4-7\alpha}{2}}\right]$$

$$\| y - Y\|_{\hat{L}_\infty(L_2(\Omega))} \le O\left[h^{\frac{2k - 3\alpha qk - 4\alpha q + 2}{2}}\right] + O\left[\Delta t^{\frac{4-7\alpha}{2}}\right] \quad \blacksquare$$

Theorem 6.9 leads to a corollary which gives a sufficient condition for the convergence of the parabolic regularization method to shock wave solutions:

Corollary 6.9. If the hypotheses of Theorem 6.7 are satisfied and $\Delta t/h^q = C$, the convergence of the parabolic regularization method (6.5.7-6.5.8) to (6.8.1) occurs if

$$k+1 > \alpha q\left(\frac{3k+6}{2}\right)$$

and

$$\alpha < \frac{4}{7} \qquad \blacksquare$$

We observe that the constraint on α given in Corollary 6.9 implies that the Lax-Wendroff type scheme (6.5.7-6.5.8) (for which $\alpha = 1$) will not necessarily converge to shock wave solutions.

The final theorem gives the criteria for numerical stability of the method.

Theorem 6.10. The parabolic regularization method (6.5.7) and (6.5.8) is numerically stable in the L_2 sense in the displacements and velocities if for arbitrarily chosen positive constants ψ, ζ

$$\frac{\Delta t^{\alpha}}{h^2} < \frac{8\rho}{C'C^{*2}}$$

$$\frac{h^2}{\Delta t^2} < \frac{C'C^{*2}}{\psi}$$

$$\frac{\Delta t}{h^2} < \frac{\psi^2}{2\lambda C^{*2}\zeta C'}$$

$$\frac{\Delta t^{1-2\alpha}}{h^2} < \frac{\psi^2\zeta}{2\lambda C^{*2}C'}$$

where $\lambda = 1 - \psi h^2/C'C^{*2}\Delta t$. $\blacksquare$

VI.9 The Nonlinear Parabolic Regularization Approximation. In this section of the chapter, we consider the approximation of (6.3.1) by (6.5.7) - (6.5.8). However, we can easily split up the second order equation (6.3.1) into two coupled first order equations by defining $y = \frac{\partial u}{\partial t}$

$$\begin{aligned} &(\rho\frac{\partial y}{\partial t},v)_o + a(u,u,v) = 0 \quad \forall \quad v \in H^1(\Omega) \\ &(\rho\frac{\partial u}{\partial t},v)_o - (y,v)_o = 0 \quad \forall \quad v \in H^1(\Omega) \end{aligned} \tag{6.9.1}$$

Thus, an equivalent problem and the problem to be undertaken here is to show the convergence of the parabolic regularization method (6.5.7) - (6.5.8) to (6.9.1).

Now we pose an auxiliary problem. Let $(\tilde{u},\tilde{y})$ be the solution to the system

$$\begin{aligned} &(\rho\frac{\partial \tilde{y}}{\partial t},v)_o + a(\tilde{u},\tilde{u},v) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{u},\tilde{y},v) = 0 \quad \forall\ v \in H^1(\Omega) \\ &(\rho\frac{\partial \tilde{u}}{\partial t},v)_o - (\rho\tilde{y},v)_o + \frac{\Delta t^{*\alpha}}{2} a(\tilde{u},\tilde{u},v) = 0 \quad \forall\ v \in H^1(\Omega) \end{aligned} \tag{6.9.2}$$

We obtain an approximate auxiliary problem by introducing the forward difference operator in (6.9.2). Then the solution to the approximate auxiliary problem $(\tilde{U}^n,\tilde{Y}^n)$ satisfies

$$\begin{aligned} &(\rho\frac{\tilde{Y}^{n+1} - \tilde{Y}^n}{\Delta t},V) + a(\tilde{U}^n,\tilde{U}^n,V) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{U}^n,\tilde{Y}^n,V) = 0 \quad \forall \quad V \in S_h(\Omega) \\ &(\rho\frac{\tilde{U}^{n+1} - \tilde{U}^n}{\Delta t},V) - (\rho\tilde{Y}^n,V) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{U}^n,\tilde{U}^n,V) = 0 \quad \forall \quad V \in S_h(\Omega) \end{aligned} \tag{6.9.3}$$

We follow here a procedure identical to that of Section 6.8. That is, we use the convergence of the auxiliary problem as an intermediate step in the proof of convergence of (6.5.7-6.5.8) to (6.9.1).

If $(6.9.2)_2$ is evaluated at time point $t = n\Delta t$ and v is equated to V, then it can be seen that

$$(\rho\frac{\partial\tilde{u}_n}{\partial t},V)_o - (\rho\tilde{y}_n,V)_o + \frac{\Delta t^{*\alpha}}{2}\ a(\tilde{u}_n,\tilde{u}_n,V) = 0 \quad \forall\ V \in S_h(\Omega) \tag{6.9.4}$$

Now adding $(\tilde{u}_{n+1} - \tilde{u}_n/\Delta t,V)_o$ to each side of (6.9.4) gives

$$(\rho\frac{\tilde{u}_{n+1} - \tilde{u}_n}{\Delta t},V)_o - (\rho\tilde{y}_n,V)_o + \frac{\Delta t^{*\alpha}}{2} a(\tilde{u}_n,\tilde{u}_n,V) = (\rho\psi_n,V)_o$$

$$\forall\ V \in S_h(\Omega) \tag{6.9.5}$$

where ψ_n is defined in (6.8.5).

Now we set $e_n = \tilde{u} - \tilde{U}^n$ and $f_n = \tilde{y} - \tilde{Y}^n$. Then subtracting $(6.9.3)_2$ from (6.9.5)

$$(\rho\frac{e_{n+1} - e_n}{\Delta t},V)_o - (\rho f_n,V)_o + \frac{\Delta t^{*\alpha}}{2}[a(\tilde{u}_n,\tilde{u}_n,V) - a(\tilde{U}_n,\tilde{U}_n,V)] = (\rho\psi_n,V)_o$$

$$\forall\ V \in S_h(\Omega) \tag{6.9.6}$$

Now we identify W^n, $P^n \in S_h(\Omega)$ through the weighted $H^1(\Omega)$ projections

$$a(\tilde{u}_n,\tilde{u}_n - W^n,V) = 0 \quad \forall\ V \in S_h(\Omega)$$

$$a(\tilde{u}_n,\tilde{y}_n - P^n,V) = 0 \quad \forall\ V \in S_h(\Omega)$$

Then we perform the normal decomposition of the error e_n and f_n. $e_n = E_n + \mathcal{E}_n$ where $E_n = \tilde{u}_n - W^n$ and $\mathcal{E}_n = W^n - \tilde{U}^n$ and $f_n = F_n + \mathcal{F}_n$ where $F_n = \tilde{y}_n - P^n$ and $\mathcal{F}_n = P^n - \tilde{Y}^n$.

Then using a method of proof quite similar to the one presented in Theorem 6.4, we have

<u>Theorem 6.11</u>. Let $\frac{\partial^3 \tilde{u}}{\partial t^3} \in L_2(L_2(\Omega))$, then there exists a positive constant C_1 such that

$$\|\mathcal{E}\|_{\hat{L}_\infty(L_2(\Omega))} \leq C_1\{\|\mathcal{E}_0\|_0 + \|\mathcal{F}\|_{\hat{L}_\infty(L_2(\Omega))} + \|\frac{\partial E}{\partial t}\|_{L_2(L_2(\Omega))}$$

$$+ \|F\|_{\hat{L}_\infty(L_2(\Omega))} + \Delta t^2 \|\frac{\partial^3 \tilde{u}}{\partial t^3}\|_{L_2(L_2(\Omega))}\} \tag{6.9.7}$$

■

Now we estimate the error induced in the approximation of $(6.9.1)_1$ by $(6.9.2)_1$. If $(6.9.1)_1$ is evaluated at $t = n\Delta t$ and v is set equal to $V \in S_h(\Omega)$, then

$$(\rho\frac{\partial \tilde{y}_n}{\partial t}, V)_0 + a(\tilde{u}_n, \tilde{u}_n, V) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{u}_n, \tilde{y}_n, V) = 0 \quad \forall \quad V \in S_h(\Omega) \tag{6.9.8}$$

Now adding $(\tilde{y}_{n+1} - \tilde{y}_n/\Delta t, V)_0$ to each side of (6.9.8), we get

$$(\frac{\tilde{y}_{n+1} - \tilde{y}_n}{\Delta t}, V)_0 + a(\tilde{u}_n, \tilde{u}_n, V) + \frac{\Delta t^{*\alpha}}{2} a(\tilde{u}_n, \tilde{y}_n, V) = (\rho\beta_n, V)_0$$

$$\forall \quad V \in S_h(\Omega) \tag{6.9.9}$$

where β_n is defined in (6.8.18).

Now subtracting $(6.9.3)_1$ from (6.9.9)

$$(\rho\frac{f_{n+1}-f_n}{\Delta t},V)_0 + a(\tilde{u}_n,\tilde{u}_n,V) - a(\tilde{U}_n,\tilde{U}_n,V)$$

$$+\frac{\Delta t^{*\alpha}}{2}[a(\tilde{u}_n,\tilde{y}_n,V) - a(\tilde{U}_n,\tilde{Y}_n,V)] = (\rho\beta_n,V)_0 \qquad (6.9.10)$$

The behavior of the approximation error F_n is established using a technique similar to the one used to prove Theorem 6.5. We state the result here

Theorem 6.12. Suppose $\frac{\partial^3 y}{\partial t^3} \in L_2(L_2(\Omega))$ and suppose that the stability condition (6.8.21) is satisfied. Then there exists a positive constant C_2 such that

$$||F||_{\hat{L}_\infty(L_2(\Omega))} \leq C_2\{||F_0||_0 + ||E||_{L_2(L_2(\Omega))} + ||\frac{\partial F}{\partial t}||_{L_2(L_2(\Omega))}$$

$$+ \Delta t^2 ||\frac{\partial^3\tilde{y}}{\partial t^3}||_{L_2(L_2(\Omega))}\} \qquad (6.9.11)$$

■

We obtain a final estimate for the rate of convergence and a sufficient condition for convergence which are the same as Theorem 6.9 and Corollary 6.9, respectively. They will not be repeated here.

CHAPTER VII

FINITE ELEMENT IMPLEMENTATION OF THE PARABOLIC REGULARIZATION PROCEDURE

VII.1 Introduction. In this chapter the parabolic regularization method introduced in Chapter VI is implemented using a particular finite element subspace, and numerical results are presented for wave propagation in one-dimensional hyperelastic bodies. In particular, we attempt to verify certain theoretical results obtained in Chapter VI In Chapter VI, we showed that the regularization parameter α must be restricted to a certain interval in order to obtain convergence to shock wave solutions. In this chapter we will attempt to experimentally verify this conclusion. In addition, we present numerical results for the propagation and reflection of shock waves.

VII.2 The Finite Element Model. Initially we define a one-dimensional domain $I = [0,L_0]$. We discretize I by defining I as the union of N-1 open sets of length $h = L_0/(N-1)$. h represents the mesh parameter.

In order to obtain a finite element model the following set of basic functions are defined on I:

$$\Phi_1(X) = 1 - \frac{X}{h}\ , \qquad 0 \le X \le h$$

$$\Phi_\alpha(X) = \begin{cases} \frac{X}{h} - (\alpha-2) & (\alpha-2)\, h \le X \le (\alpha-1)h \\ \alpha - \frac{X}{h} & (\alpha-1)\, h \le X \le \alpha h \qquad (\alpha=2,\ldots,N-1) \end{cases}$$

$$\Phi_N(X) = \frac{X}{h} - (N-2) \qquad (N-2)h \le X \le (N-1)h \tag{7.2.1}$$

We define one-dimensional finite element models for the displacement and velocity fields in terms of these basis functions by

$$U = \sum_{\alpha=1}^{N} U_\alpha \Phi_\alpha$$

$$V = \sum_{\alpha=1}^{N} V_\alpha \Phi_\alpha$$

These representations are piecewise linear. This implies that both the gradient of U and the gradient of V are constant across an element.

We now construct the finite-element/parabolic regularization equations for the one-dimensional elastic rod. We use here the formulation introduced in Chapter II. Initially, we define the quasi-bilinear forms $a(U^n,U^n,V)$ and $a(U^n,Y^n,V)$ appearing in (6.5.7) and (6.5.8) by

$$a(U^n,\ U^n,\ V) = \int_I \sigma(U_X^n) V_X dX \tag{7.2.3}$$

and

$$a(U^n,\ Y^n,\ V) = \int_I \frac{\partial\sigma(U_X^n)}{\partial U_X^n}\ Y_X^n V_X dX \tag{7.2.4}$$

Then introducing (7.2.2) into (6.5.7) and (6.5.8) and using (7.2.3) and (7.2.4), we obtain the finite element/parabolic regularization equations.

$$\frac{1}{\Delta t} M_{\alpha\beta} U_\beta^{N+1} - \frac{1}{\Delta t} M_{\alpha\beta} U_\beta^n - M_{\alpha\beta} Y_\beta^n$$

$$+ \frac{\Delta t^\alpha}{2} \int_I \sigma(U_X^n) \Phi_{\alpha,X} dX = \frac{\Delta t^\alpha}{2} f_\alpha \tag{7.2.5}$$

$$\frac{1}{\Delta t} M_{\alpha\beta} Y_\beta^{n+1} - \frac{1}{\Delta t} M_{\alpha\beta} Y_\beta^{n} + \frac{\Delta t^\alpha}{2} \int_I \frac{\partial\sigma(U_X^n)}{\partial U_X^n} Y_X^n \Phi_{\alpha,X} dX$$

$$+ \int_I \sigma(U_X^n)\Phi_{\alpha,X} dX = f_\alpha + \frac{\Delta t^\alpha}{2} P_\alpha \qquad (7.2.6)$$

where

$$M_{\alpha\beta} = \int_I \Phi_\alpha \Phi_\beta dX$$

$$f_\alpha = \int_I f\Phi_\alpha dX$$

$$P_\alpha = \int_I \frac{\partial f}{\partial t} \Phi_\alpha dX$$

In conjunction with (7.2.5) and (7.2.6) initial values of U and Y are defined by

$$\int_I U^o \, \Phi_\alpha dX = \int_I u(0)\Phi_\alpha dX$$

$$\qquad (7.2.7)$$

$$\int_I Y^o \, \Phi_\alpha dX = \int_I \frac{\partial u}{\partial t}(0)\Phi_\alpha dX$$

VII.3 Numerical Results. Initially the case of linear wave propagation is considered. We have selected here for calculation the simple example

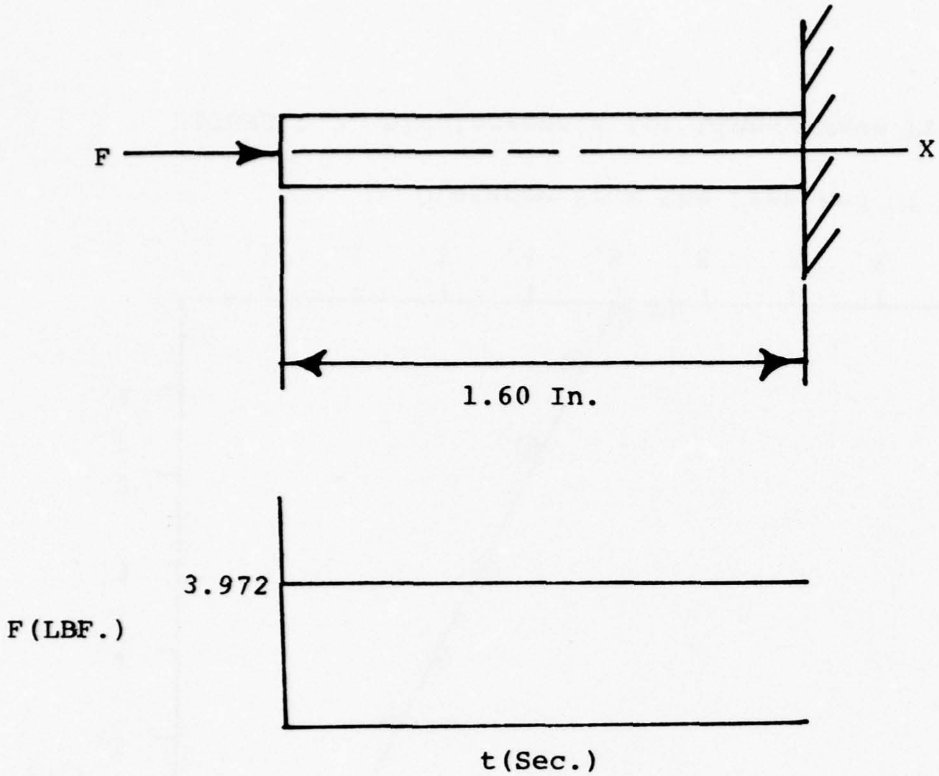
of wave propagation in a one-dimensional linear elastic bar with step load at the free end. The finite element model and physical parameters are given in Figure 7.1.

The goals of this work are to show that the finite element model converges for linear wave propagation, to establish the ranges of the regularization parameter α for which convergence can be obtained, and to compare these results with the theoretical range of convergence for α. The initial studies were performed with $\alpha = 2.0$. In figures .2 and .3 the displacement approximation is compared to the exact solution with the wave positioned approximately one-third of the distance down the bar. The results in Figure 7.2 were obtained with the discretization No. 1 (a coarse mesh). The results in Figure 7.3 were obtained with discretization No. 2 (a finer mesh). The two discretizations have

$$\frac{\Delta t}{h} = C$$

The displacement approximation produced of discretization No. 2 is clearly more accurate than the one produced by discretization No. 1. Thus for $\alpha = 2.0$, the displacement approximation converges. In Figure 7.4 and 7.5 the velocity approximation is compared to the exact solution for the same two discretizations. Clearly the approximation for velocity of discretization No. 2 is worse than that produced by discretization No. 1. This indicates that for $\alpha = 2.0$ the velocities do not converge. Thus for $\alpha = 2.0$ the parabolic regularization method does not converge for linear wave propagation.

More reasonable results can be obtained with $\alpha = 0.8$. In Figures 7.6 and 7.7 the displacement approximation associated with



PHYSICAL PARAMETERS

Density = 1.0×10^{-4} $\frac{\text{LBF.} - \text{Sec.}^2}{\text{In.}^4}$

Area = .0314 In.^2

Modulus of Elasticity = 63.25 LBF./In.^2

Length = 1.60 In.

FIGURE 7.1. Physical Model for One-Dimensional Linear Wave Propagation

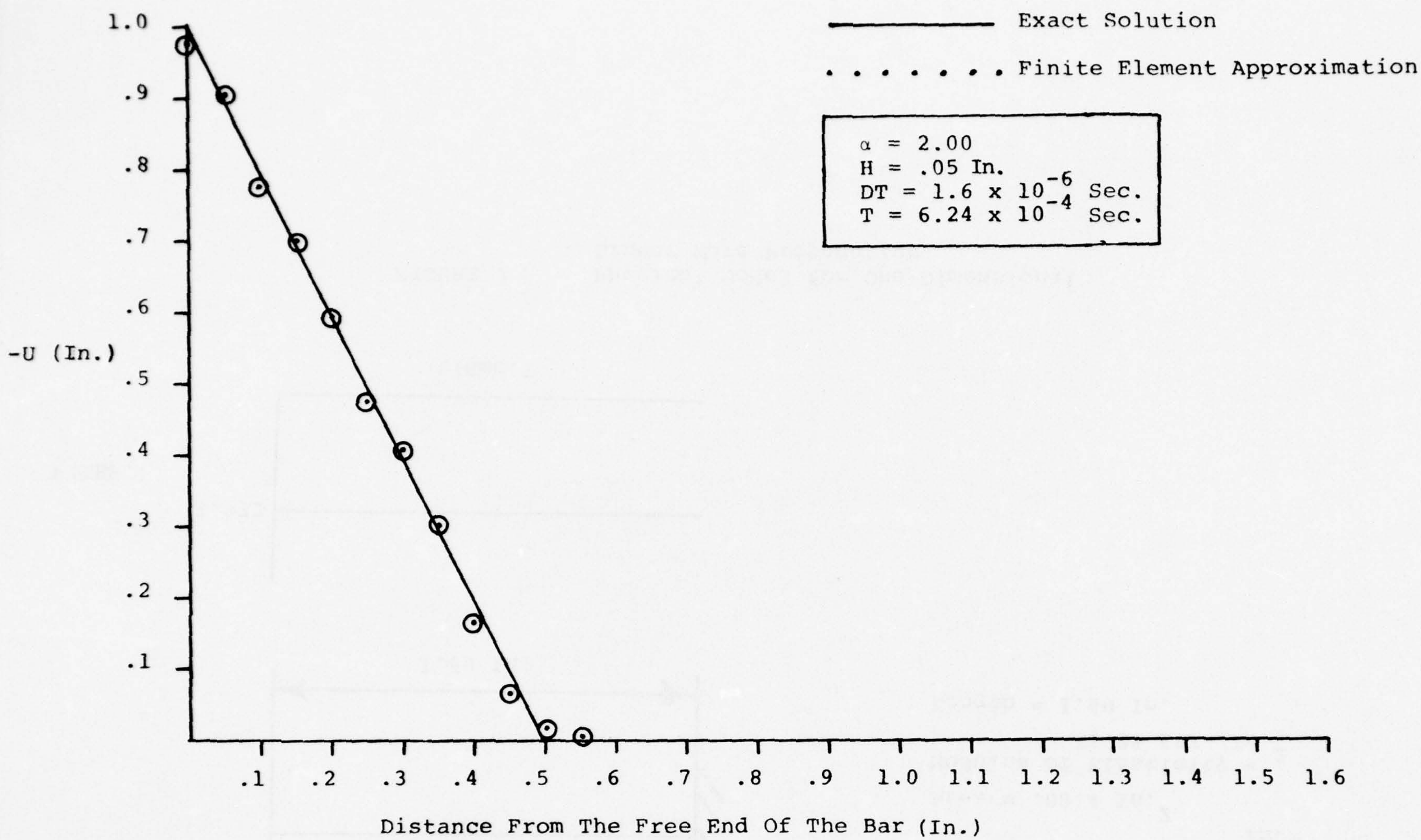


FIGURE 7.2 Displacements for Linear Wave Propagation with α = 2.0 - Discretization No. 1

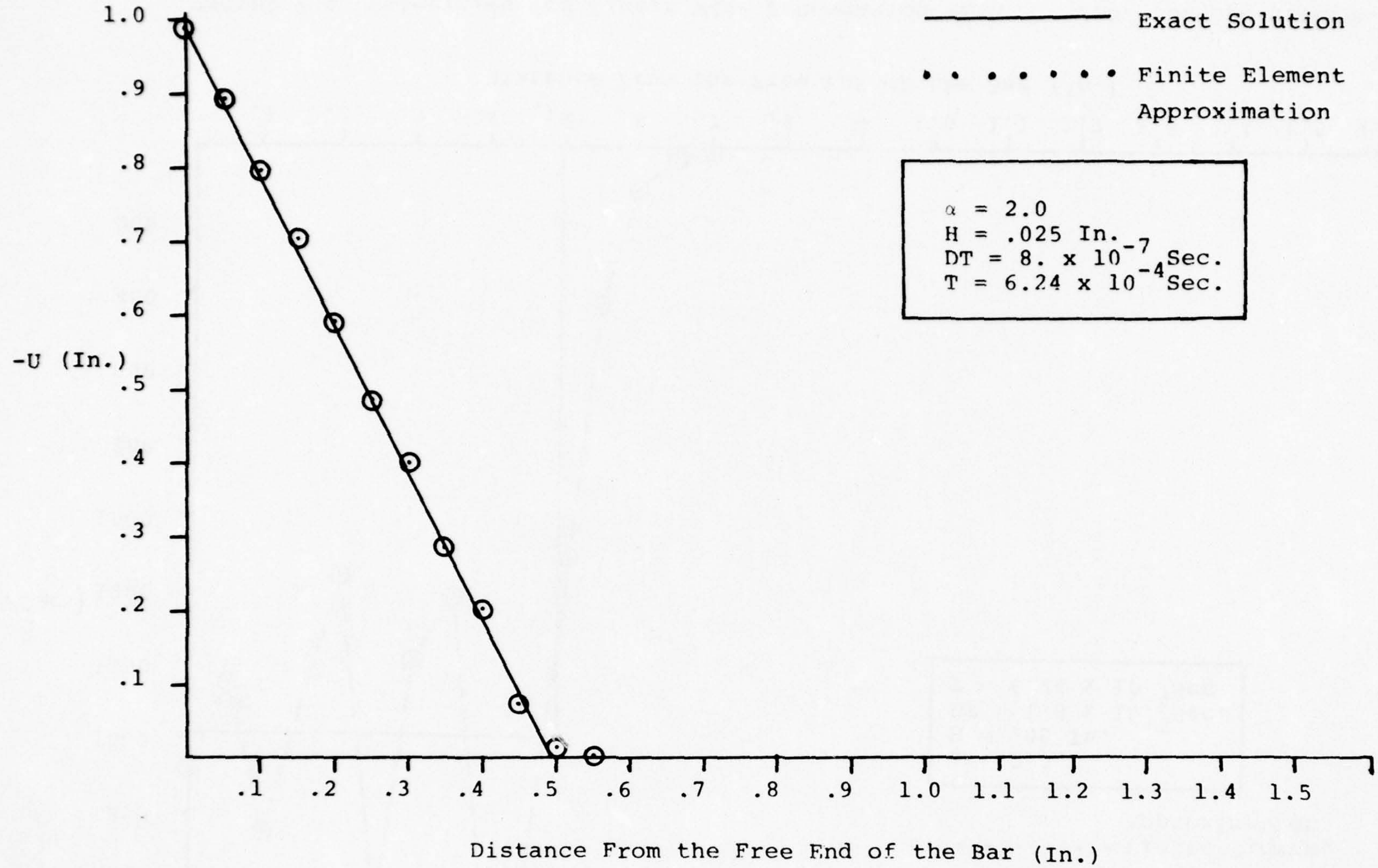


FIGURE 7.3 Displacements for Linear Wave Propagation with $\alpha = 2.0$ - Discretization No. 2

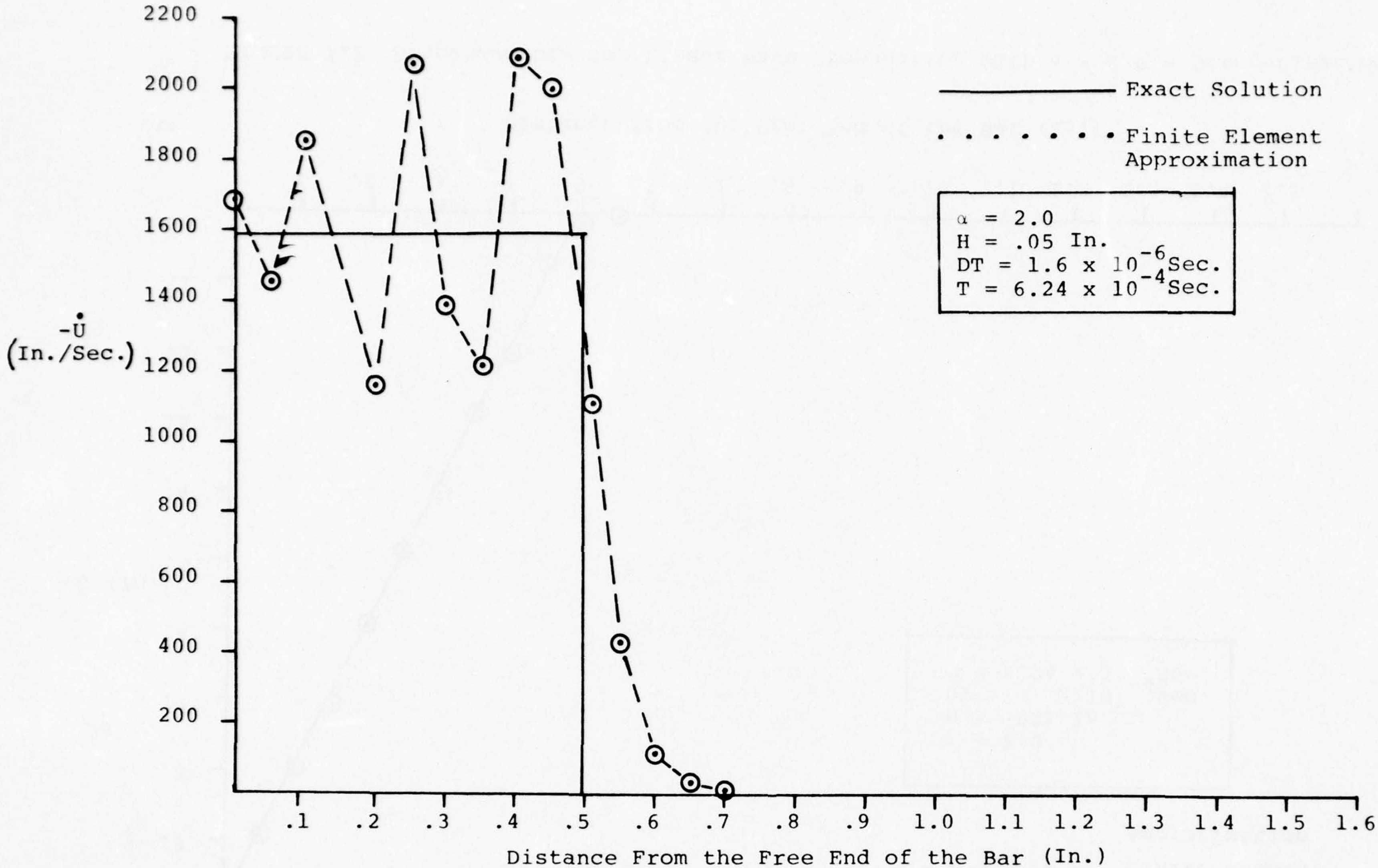


FIGURE 7.4 Velocities for Linear Wave Propagation With $\alpha = 2.0$ - Discretization No. 1

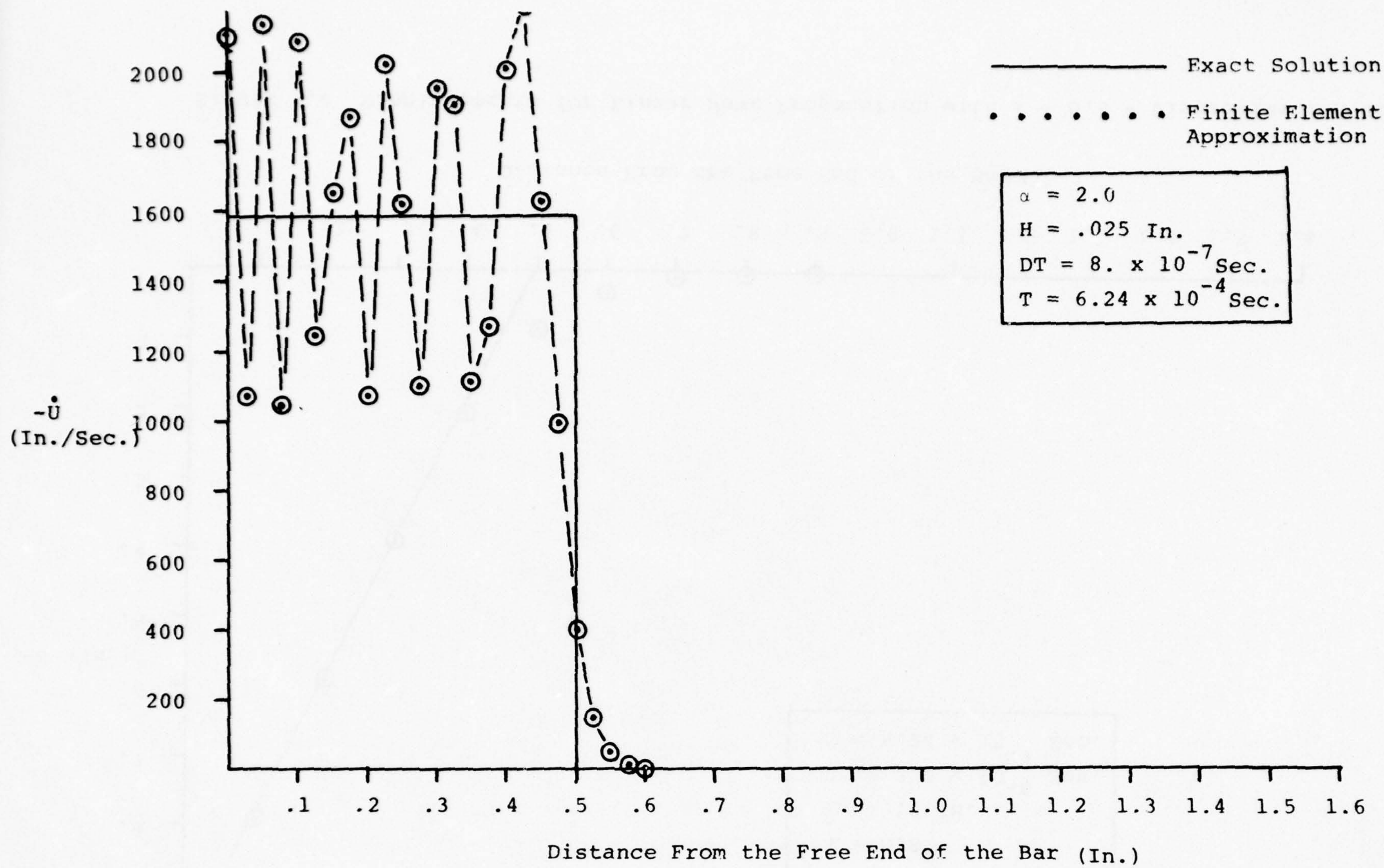


FIGURE 7.5 Velocities for Linear Wave Propagation with α = 2.0 - Discretization No. 2

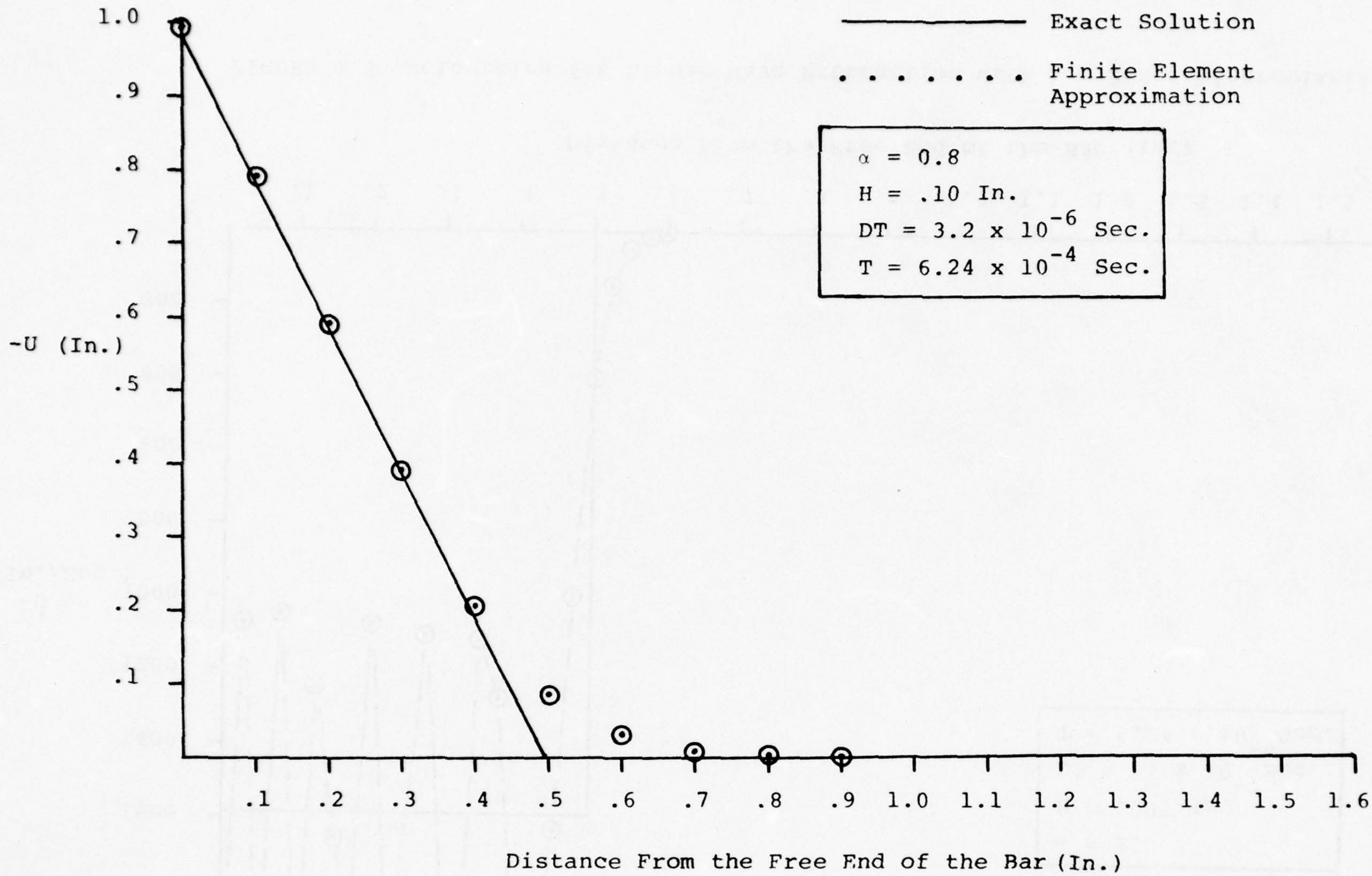


FIGURE 7.6 Displacements for Linear Wave Propagation with $\alpha = 0.8$ - Discretization No. 1

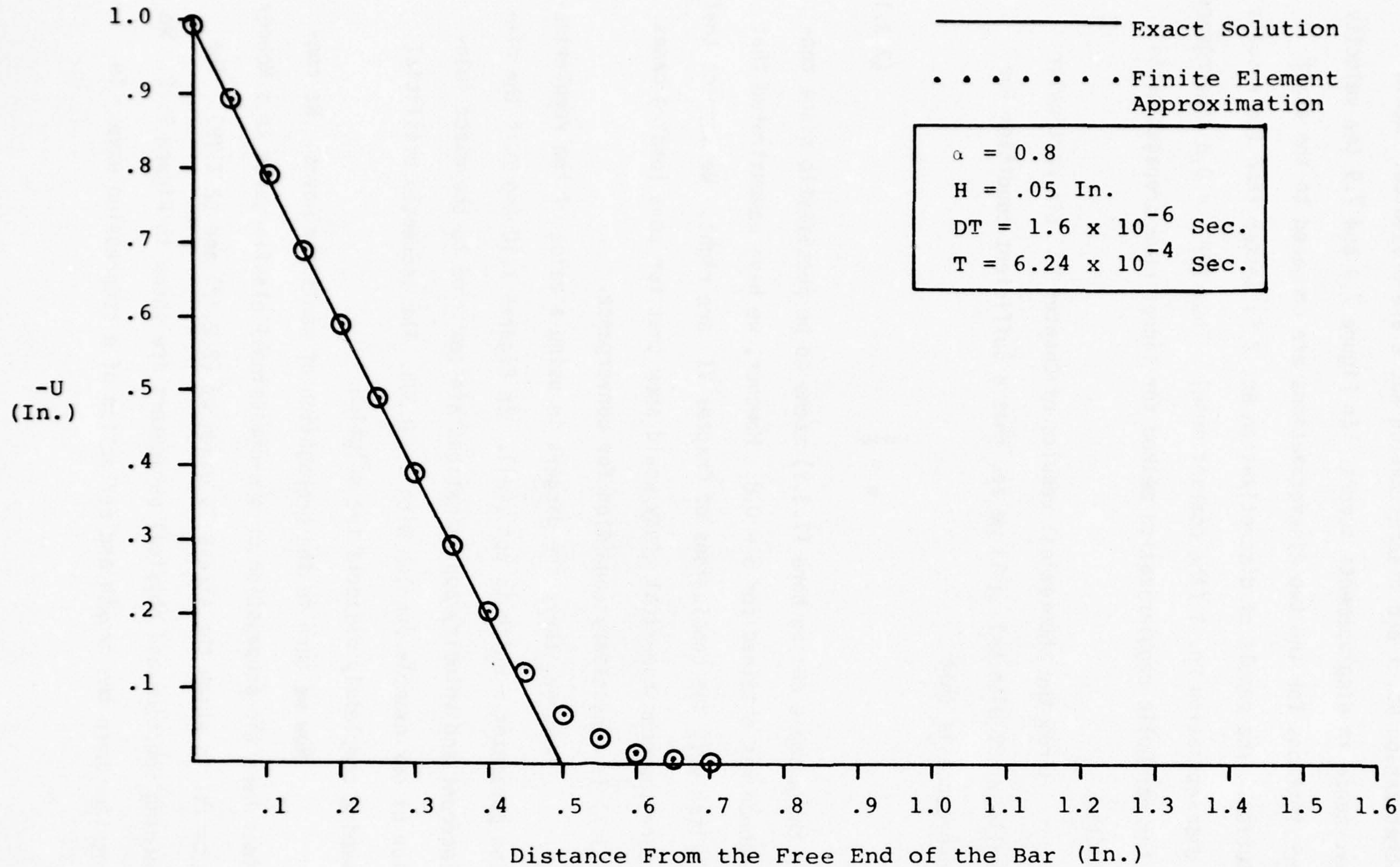


FIGURE 7.7 Displacements for Linear Wave Propagation with α =0.8 - Discretization No. 2

discretization No. 1 and discretization No. 2 are presented. Again convergence in displacements occurs. In Figure 7.8 and 7.9 the velocity approximations for the two discretizations are compared to the exact solution. The result of discretization No. 2 is better than the results of discretization No. 1 (the coarser mesh). Thus for $\alpha = 0.8$ convergence of the parabolic regularization method for linear wave propagation is obtained.

From the theoretical results of Chapter VI (in particular Corollary 6.9 with k=1, q=1) we see that a sufficient condition for convergence is that

$$\alpha < \frac{4}{9} \tag{7.3.1}$$

For the example chosen here (7.3.1) seems to be pessimistic since convergence was obtained for $\alpha = 0.8$. However, we have demonstrated that qualitatively the conclusions of Chapter VI are right. We suspect that a more thorough numerical study would show that for some loading cases (7.3.1) is a necessary condition for convergence.

However, there are dangers in using a value of the regularization parameter α which is too small. In Figures 7.10 and 7.11 the displacement and velocity approximations are compared to the exact solution of our example problem with $\alpha = 0.333$. The increased artificial damping completely distorts the solution.

Now we turn to the propagation of nonlinear waves. We consider here the propagation of one-dimensional elastic waves in a Mooney material in which the stress is given by (2.2.12) and (2.2.18). The forcing function and physical parameters are shown in Figure 7.12. We consider here the growth and reflection of a compression wave. In

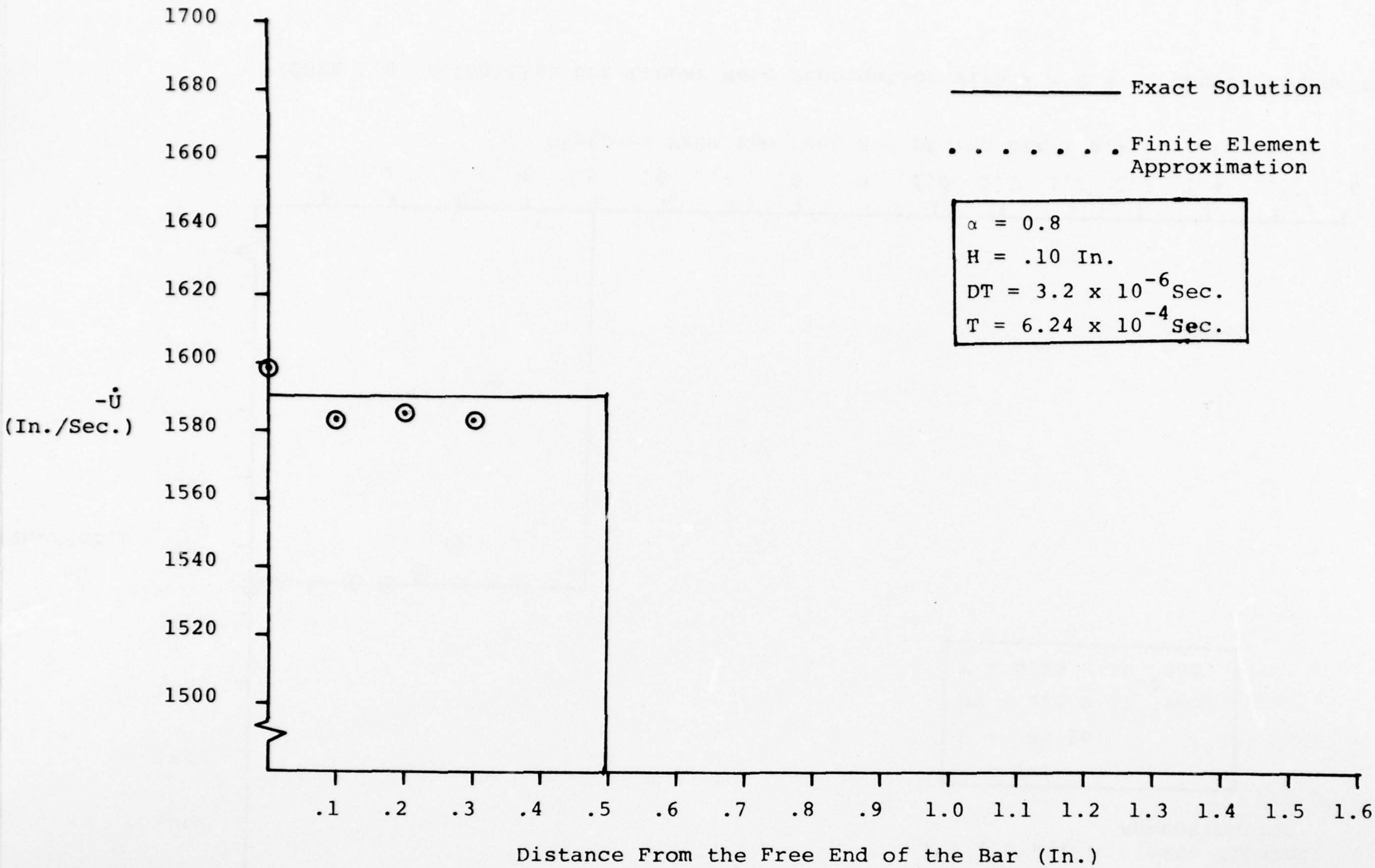


FIGURE 7.8 Velocities for Linear Wave Propagation with α = 0.8 - Discretization No. 1

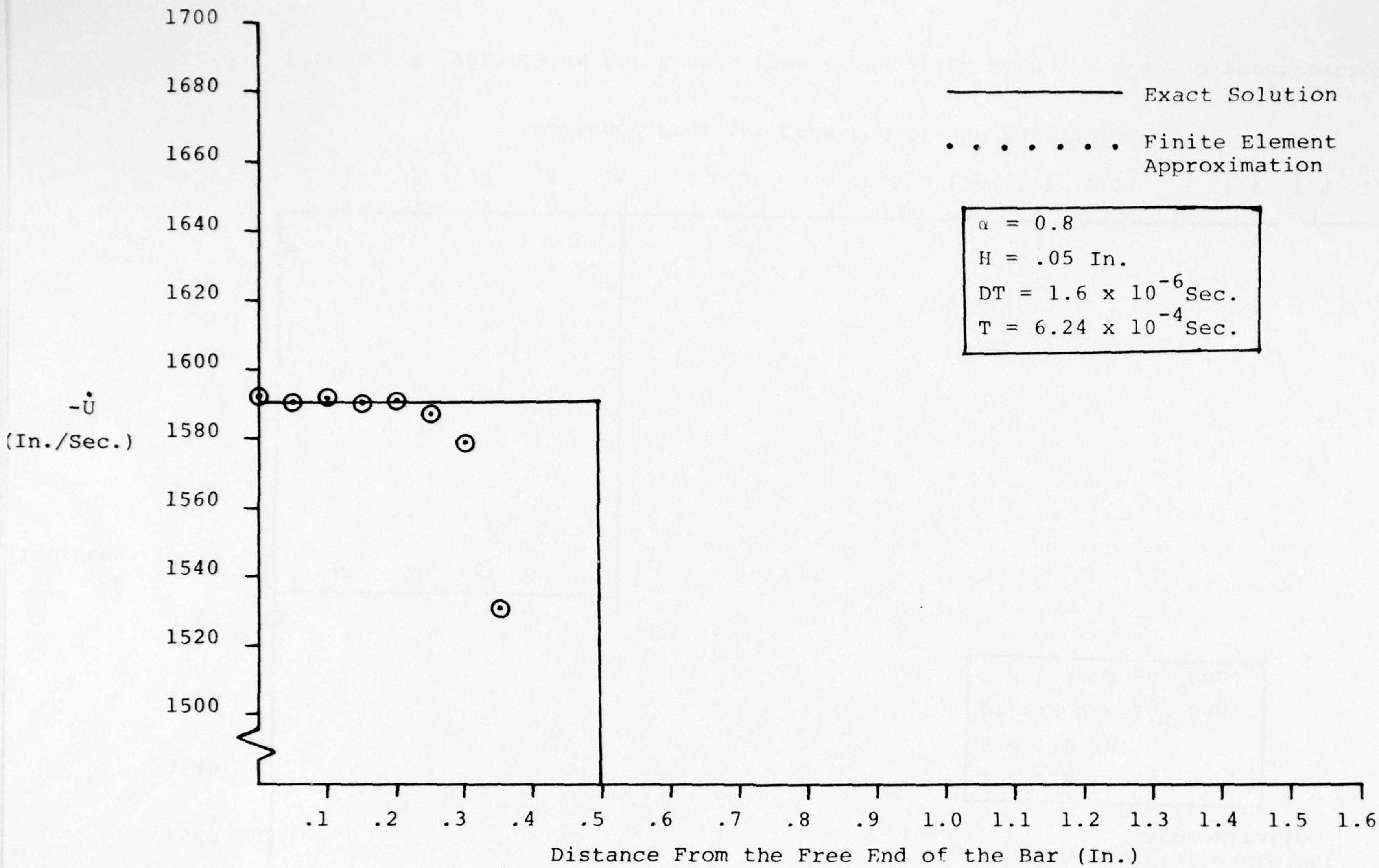


FIGURE 7.9 Velocities for Linear Wave Propagation with $\alpha = 0.8$ - Discretization No. 2

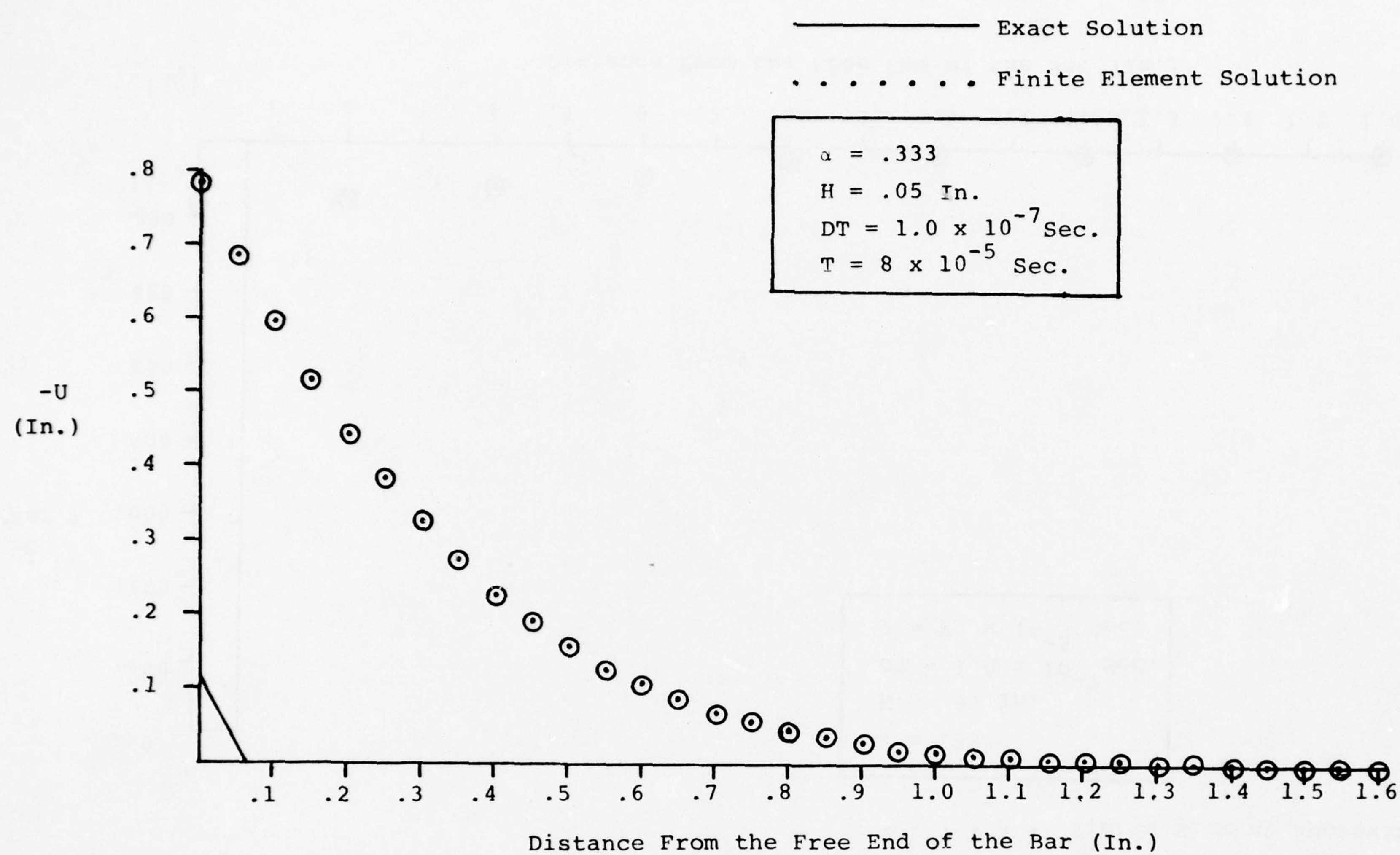


FIGURE 7.10 Displacements for Linear Wave Propagation with $\alpha = .333$

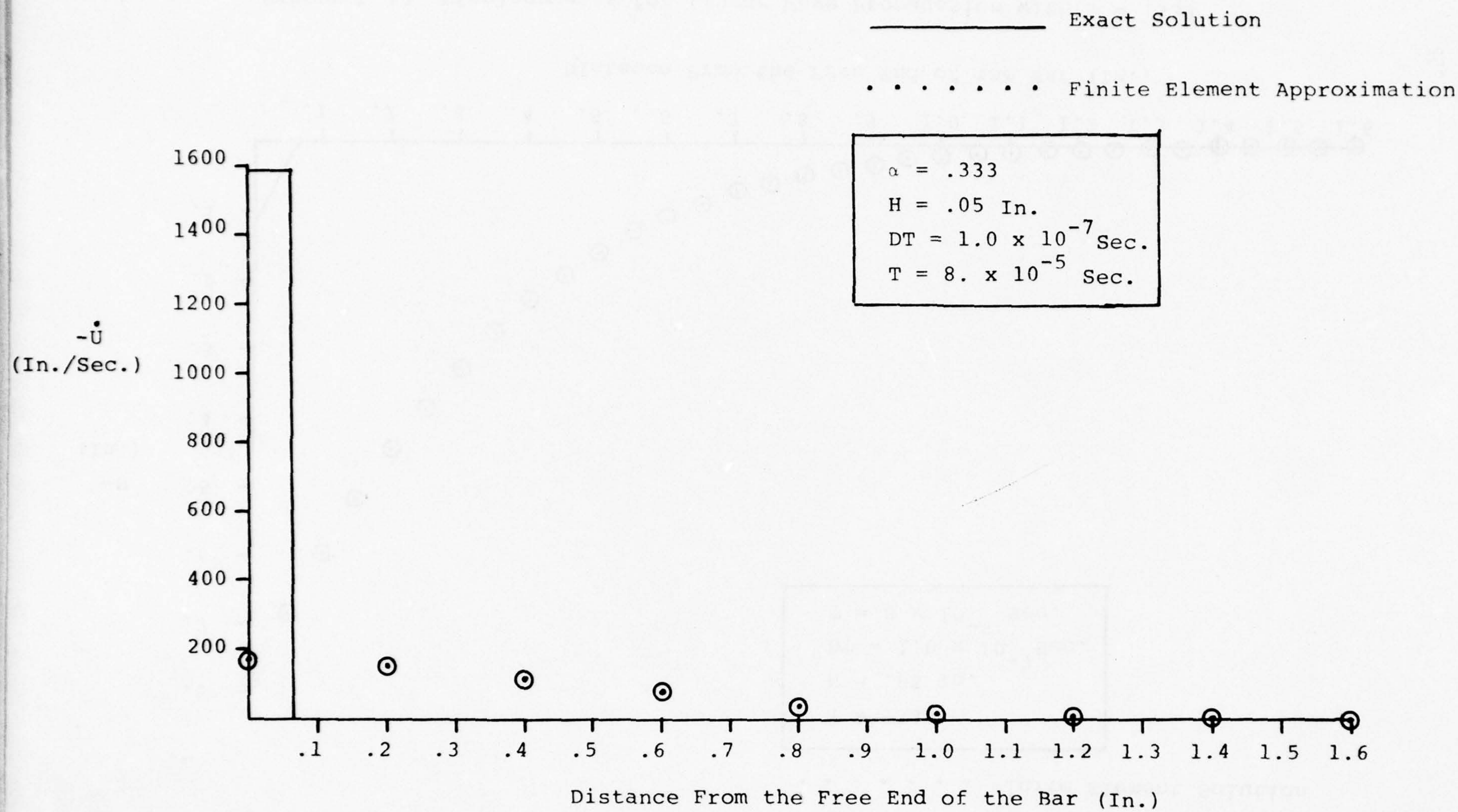
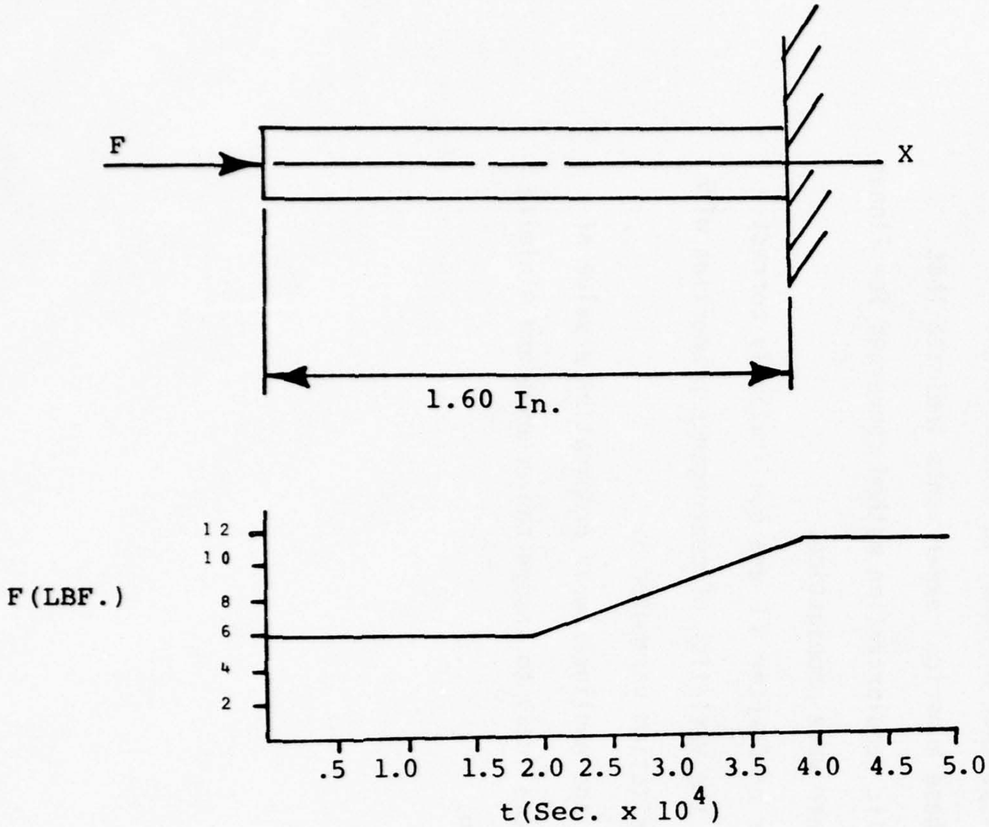


FIGURE 7.11 Velocities for Linear Wave Propagation with α = .333



PHYSICAL PARAMETERS

Density = 1.0×10^{-4} LBF.- Sec.^{2}/In.^{4}

Area = .0314 In.^{2}

C1 = 24.0 LBF./In.^{2}

C2 = 1.5 LBF./In.^{2}

Length = 1.60 In.

FIGURE 7.12 Physical Model for One-Dimensional Shock Propagation in a Mooney Material

these calculations α was set equal to 0.8. In Figure 7.13 the shape of the wave for several distinct time points is presented. In Figure 7.13 (e) the wave is shown immediately after the initial reflection. The amplitude of the stress behind the wave is increased by a factor of approximately four at the initial reflection.

The same calculation was performed with $\alpha = 0.95$. In Figure 7.14 the wave is shown at two time points. The analogous profiles for $\alpha = 0.80$ were plotted in Figure 7.13. Clearly for $\alpha = 0.95$ there is no convergence of the parabolic regularization method to shock waves. We thus suggest that for shock calculations with the parabolic regularization method a value of α less than or equal to 0.80 be used.

In conclusion these numerical experiments indicate that

(i) The parabolic regularization method converges for linear and nonlinear wave propagation.

(ii) The results of Chapter VI are qualitatively correct concerning the variation of convergence properties with the regularization parameter α.

(iii) For linear and nonlinear wave propagation a value of $\alpha = 0.8$ seems ideal to insure convergence and minimize dissipation.

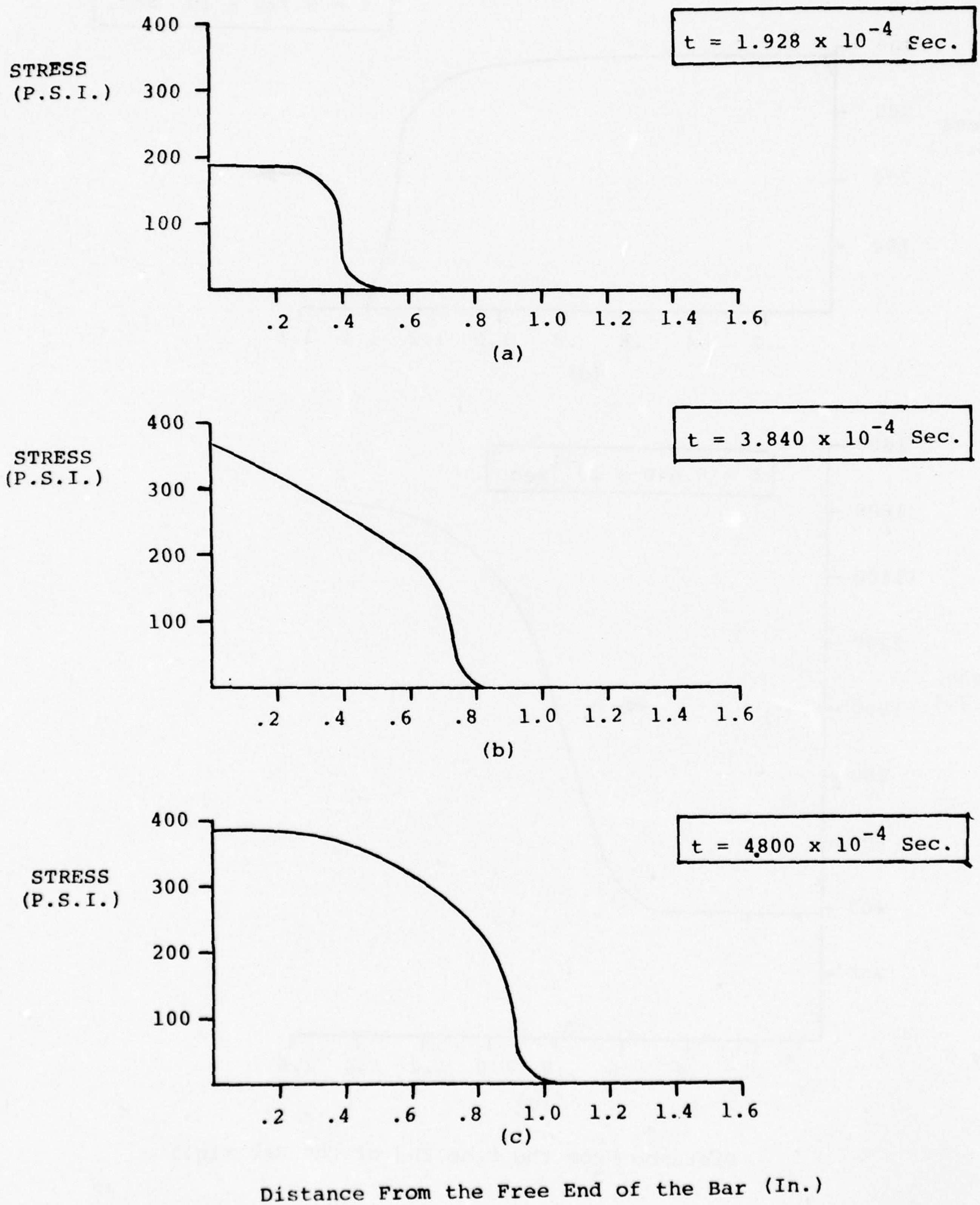
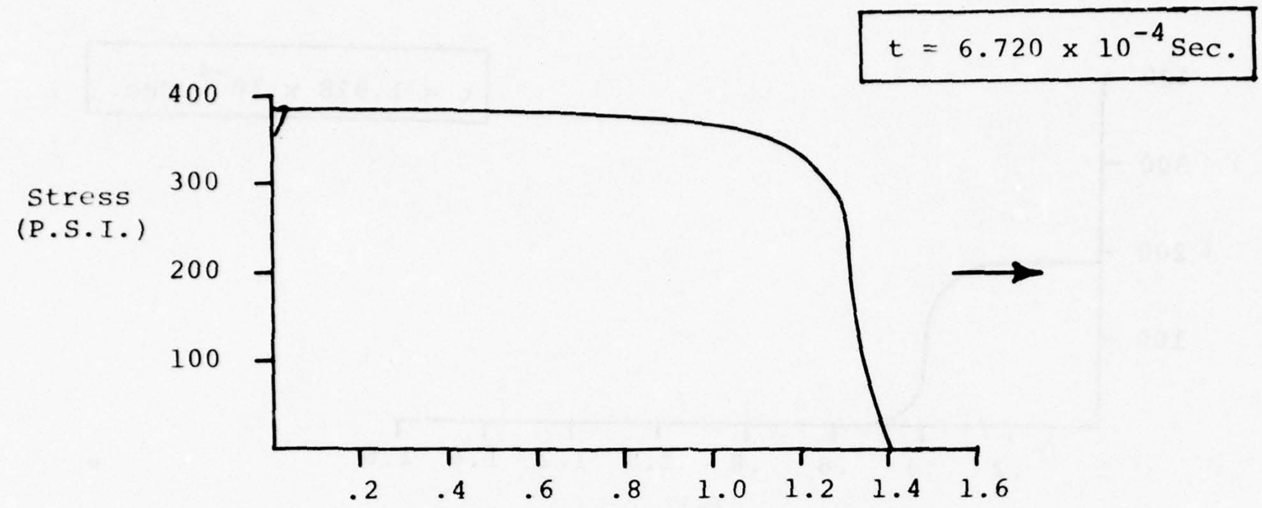
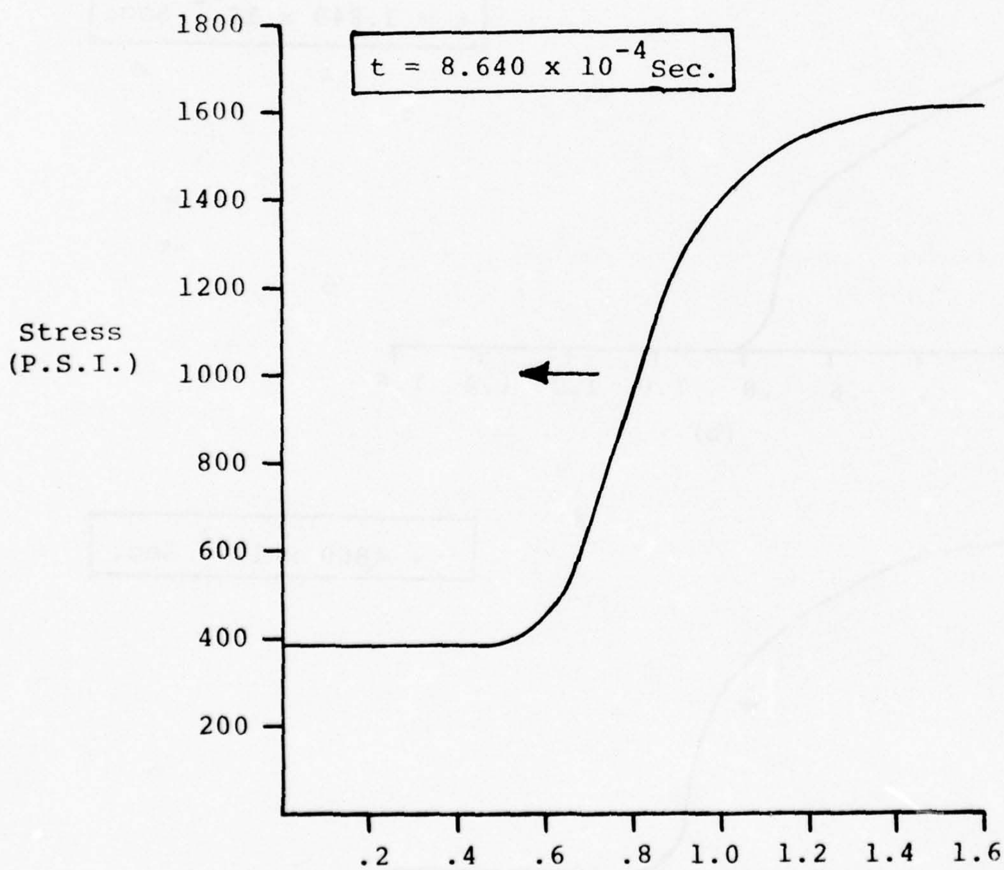


FIGURE 7.13 Compression Shock in a Mooney Material (α = .80)



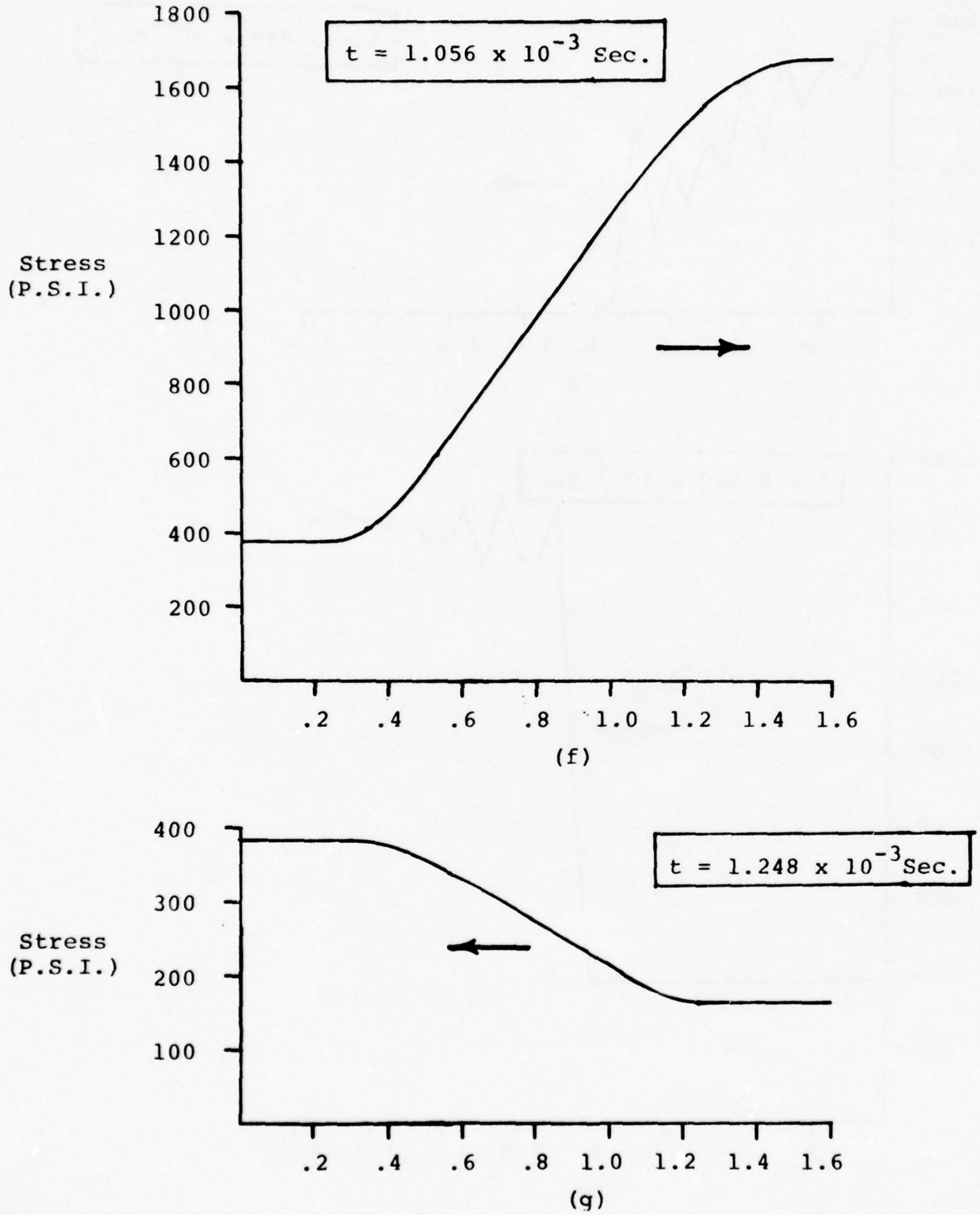
(d)



(e)

Distance From the Free End of the Bar (In.)

FIGURE 7.13 (cont.) Compression Shock in a Mooney Material (α = .80)



Distance From the Free End of the Bar (In.)

FIGURE 7.13 (cont.) Compression Shock in a Mooney Material (α = .80)

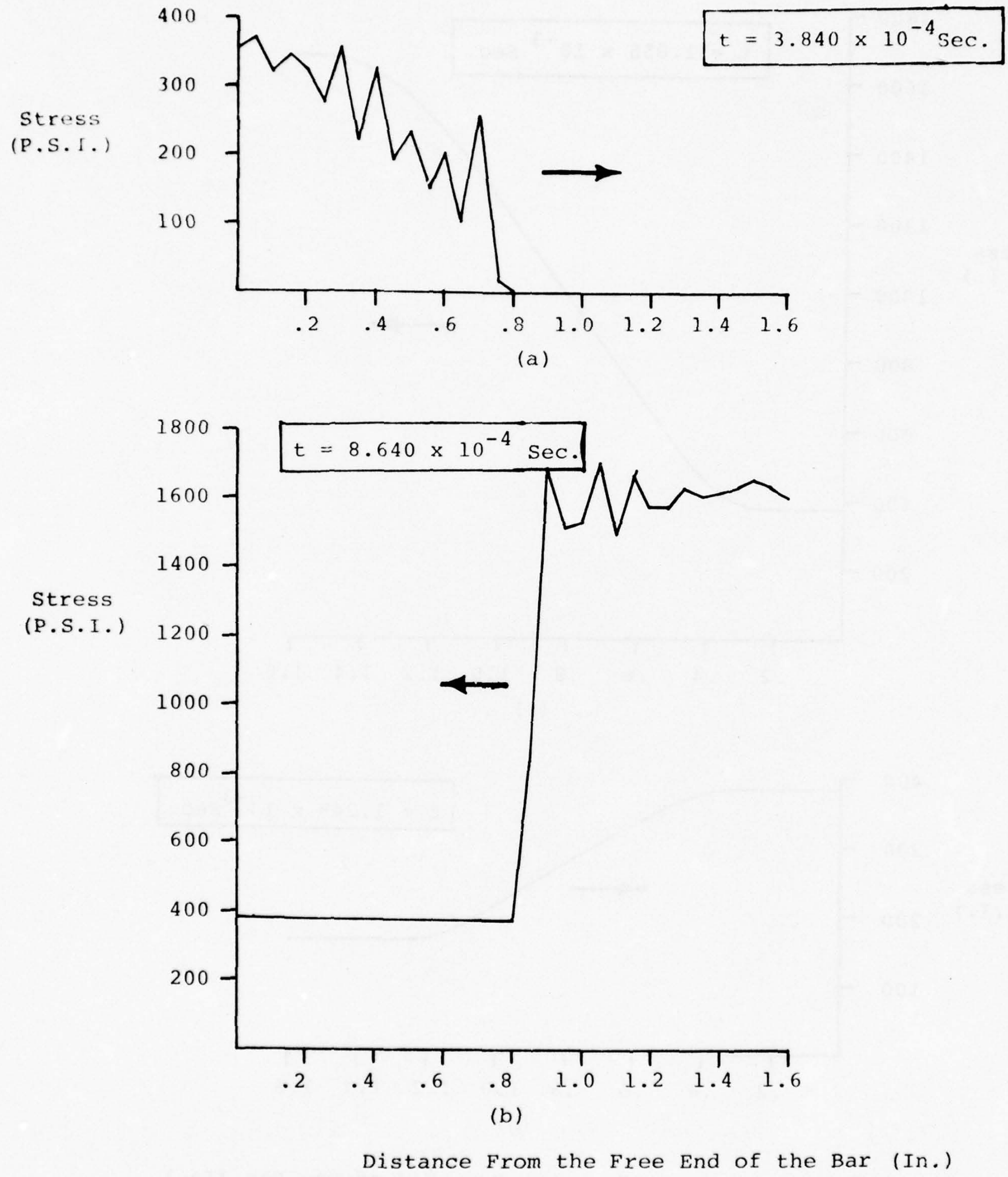


FIGURE 7.14 Compression Shock in a Mooney Material (α = .95)

CHAPTER VIII

A THEORY OF CENTRAL DIFFERENCE METHODS FOR WAVE PROPAGATION

VIII.1 Introduction. In this chapter we investigate a central difference approximation for the second order hyperbolic equation. We formulate a finite element analog for these equations and demonstrate the accuracy and convergence of the model. We determine stability conditions which turn out to be sufficient conditions for convergence.

Initially we consider the linear problem in order to demonstrate the method of deriving error estimates. We show that a version of the Courant-Friedricks-Lewy stability criteria is sufficient for stability in this case. In the nonlinear case we show that an additional condition involving the amplitude of the response must be applied to insure stability.

VIII.2 Problem Formulation. First consider a classical nonlinear hyperbolic problem of the second order characterized as follows: Find a function $u(\underset{\sim}{X},t)$, $(\underset{\sim}{X},t) \in \Omega \times [0,T]$, such that

$$\rho \frac{\partial^2 u}{\partial t^2}(\underset{\sim}{X},t) - \underset{\sim}{\nabla} \cdot (C^2(\underset{\sim}{X},t,u)\, \underset{\sim}{\nabla}\ u(\underset{\sim}{X},t)) = f(\underset{\sim}{X},t) \text{ in } \Omega \times (0,T]$$

$$u(\underset{\sim}{X},0) = g_1(\underset{\sim}{X}) \quad \text{in } \Omega$$

$$\frac{\partial u(\underset{\sim}{X},0)}{\partial t} = g_2(\underset{\sim}{X}) \quad \text{in } \Omega$$

$$u(\underset{\sim}{X},t) = 0 \qquad \text{in } \partial\Omega \times (0,T] \tag{8.2.1}$$

Even under reasonable assumptions on the functions $C^2(\underset{\sim}{X},t,u)$, $f(\underset{\sim}{X},t)$, $g_1(\underset{\sim}{X})$, and $g_2(\underset{\sim}{X})$, we often can be assured that a solution exists for this problem only for certain choices of T. For the moment, assume that such a solution exists for all $t \in (0,T]$. We remark that the notation $C^2(\underset{\sim}{X},t,u)$ is used in recognition of this function as the square of the intrinsic wave speed at which "disturbances" are propagated in problem (8.2.1).

It is well known that a solution to the problem (8.2.1) is also a solution of the following weaker problem: Find

$u(\underset{\sim}{X},t) \in L_2(H_0^1(\Omega))$, $(\underset{\sim}{X},t) \in \Omega \times [0,T]$ such that

$$\left(\rho \frac{\partial^2 u}{\partial t^2}, v\right)_0 + a(u,u,v) = (f,v)_0 \quad \forall \; v \in H^1(\Omega)$$

$$(u(\cdot,0),v)_0 = (g_1,v)_0 \quad \forall \; v \in H^1(\Omega) \tag{8.2.2}$$

$$\left(\frac{\partial u}{\partial t}(\cdot,0),v\right)_0 = (g_2,v)_0 \quad \forall \; v \in H^1(\Omega)$$

where

$$a(u,u,v) = \int_\Omega C^2(\underset{\sim}{X},t,u)\, \underset{\sim}{\nabla} u \cdot \underset{\sim}{\nabla} v \, dx \tag{8.2.3}$$

Here we complete the problem description by characterizing the real valued function $C^2(\underset{\sim}{X},t,u)$: we assume that there exist positive constants M_1, M_2, and M_3 such that

$$\text{(i)} \quad C^2(\underset{\sim}{X},t,u) \geq M_1 \quad \forall\ (\underset{\sim}{X},t) \in \Omega \times [0,T]$$

$$\text{(ii)} \quad C^2(\underset{\sim}{X},t,u) \leq M_2 \quad \forall\ (\underset{\sim}{X},t) \in \Omega \times [0,T] \tag{8.2.4}$$

$$\text{(iii)} \quad |C^2(\underset{\sim}{X},t,u) - C^2(X,t,\bar{u})| \leq M_3|u - \bar{u}| \quad \forall(\underset{\sim}{X},t) \in \Omega \times [0,T]$$

These restrictions imply positiveness, boundedness, and Lipschitz continuity respectively, of the wave speed function C^2.

Now suppose we identify a finite dimensional subspace M of $H_o^1(\Omega)$. Then the semidiscrete Galerkin approximation U of the weak solution u of (8.2.2) is that $U \in M$ such that

$$(\rho \frac{\partial^2 U}{\partial t^2}, V)_o + a(U,U,V) = (f,V)_o \quad \forall \ V \in M$$

$$(U(\cdot,0),V)_o = (g_1,V)_o \quad \forall \ V \in M \tag{8.2.5}$$

$$(\frac{\partial U}{\partial t}(\cdot,0),V)_o = (g_2,V)_o \quad \forall \ V \in M$$

To develop finite-element models of our problem, the region Ω is partitioned into a finite number E of disjoint open sets Ω_e called finite elements:

$$\Omega = \bigcup_{e=1}^{E} \bar{\Omega}_e; \ \Omega_e \bigcap \Omega_f = 0, e \neq f \tag{8.2.6}$$

Here $\bar{\Omega}_e$ is the closure of Ω_e. We let

$$h_e = \text{dia}(\bar{\Omega}_e)$$

then

$$h = \max_{1 \leq e \leq E} \{h_e\}$$

Then we define a finite element subspace $S_h(\Omega)$ associated with this finite element mesh. It shall be assumed that $S_h(\Omega)$ has the following properties (Cf. [62], [63], [76]):

(i) Let $P_j(\Omega)$ be the space of polynomials of degree j on Ω. Then there exists an integer k such that $p(\underset{\sim}{X}) \in P_j(\Omega)$ is in $S_h(\Omega)$ as long as $j \leq k$.

(ii) Let $h \to 0$ uniformly (i.e., for each refinement of the mesh let the radius ρ_e of the largest sphere than can be inscribed in Ω_e be proportional to h_e). Then there is a constant K independent of h such that

$$\underset{W \in S_h(\Omega)}{\mathrm{INF}} \| u - W \|_{H^m(\Omega)} \leq K h^{k+1-m} \|u\|_{H^{k+1}(\Omega)} \tag{8.2.7}$$

(iii) $S_h(\Omega)$ satisfies an inverse hypothesis [76] of the following form: there exists a constant C^* independent of h such that

$$\|V\|_{H^j(\Omega)} \leq C^* h^{-j} \|V\|_{L_2(\Omega)} \quad \forall \; V \in S_h(\Omega), \; j \leq k+1 \tag{8.2.8}$$

The finite-element Galerkin model is formulated by setting $M = S_h(\Omega)$ and letting $V = \Phi_N$, $N = 1, \ldots, N_o$ in (8.2.5)

$$\left.\begin{aligned} &(\rho \frac{\partial^2 U}{\partial t^2}, \Phi_N)_o + a(U,U,\Phi_N) = (f,\Phi_N)_o \\ &(U(\cdot,0),\Phi_N)_o = (g_1,\Phi_N)_o \\ &(\frac{\partial U}{\partial t}(\cdot,0),\Phi_N)_o = (g_2,\Phi_N)_o \end{aligned}\right\} \quad \begin{aligned} &\Phi_N \in S_h(\Omega), \\ &N = 1,\ldots,N_o \end{aligned} \tag{8.2.9}$$

Let P be a partition of the time domain [0,T] of the form $\{t_0, t_1, \ldots, t_N\}$ where $0 \leq t_0 < t_1 < \ldots < t_N = T$ and $t_{n+1} - t_n = \Delta t$ for $0 \leq n \leq N-1$. The values of the dependent variable U(t) at the points of the partition P are denoted by $\{U^n\}_{n=0}^N$. In order to construct a difference approximation for (8.2.9), we introduce the central difference operator

$$\delta_t^2 U^n = \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} \tag{8.2.10}$$

The <u>nonlinear central difference approximation problem</u> corresponding to (8.2.5) is to find the sequence $\{U^n\}_{n=0}^N$, where $U^n \in S_h(\Omega)$, such that

$$(\rho\delta_t^2 U^n, V)_0 + a(U^n, U^n, V) = (f, V)_0 \quad \forall \quad V \in S_h(\Omega) \tag{8.2.11}$$

We must, of course, add appropriate approximations of initial conditions. For example,

$$(U^0, V)_0 = (g_1, V)_0, \; (\delta_t U^{1/2}, V)_0 = (g_2, V)_0 \quad \forall \quad V \in S_h(\Omega)$$

where $\delta_t U^{1/2}$ represents a forward difference operator $\dfrac{U^1 - U^0}{\Delta t}$.

<u>VIII.3.The Linear Central Difference Approximation</u>. In this section we briefly outline the method of establishing a priori bounds for the approximation error involved in modeling the linearized version of (8.2.2), with the linearized scheme (8.2.11). The methods used in this section are of the L_2-type, and represent an extension of the results of Dupont [49] to the explicit case. Our results are in some ways similar to those obtained by Fujii [77]. In the section following this one, we expand our results to the nonlinear problem.

Initially we evaluate (8.2.2) at $t = n\Delta t$ and set $v = V \in S_h(\Omega)$. Then adding $(\rho\delta_t^2 u_n, V)_o$ to each side of (8.2.2), we have

$$(\rho\delta_t^2 u_n, V)_o + a(u_n, V) = (f, V)_o + (\varepsilon_n, V)_o \quad \forall \; V \in S_h(\Omega) \tag{8.3.1}$$

where u_n is the exact solution evaluated at time point $t = n\Delta t$,

$$a(u_n, V) = \int_\Omega C^2(\underset{\sim}{x}, t)\underset{\sim}{\nabla} u_n \cdot \underset{\sim}{\nabla} V \, dx, \text{ and}$$

$$\varepsilon_n \equiv \delta_t^2 u_n - \left.\frac{\partial^2 u}{\partial t^2}\right|_{t = n\Delta t} \tag{8.3.2}$$

We assume in this development that the regularity property, $\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$ holds. This assumption precludes the existence of certain physical phenomena such as shock and acceleration waves in the solution. For $\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$, Dupont [49] has shown that an estimate for ε_n is

$$\|\varepsilon_n\|^2_{L_2(\Omega)} \leq C\Delta t^3 \int_{t_{n-1}}^{t_{n+1}} \left\| \frac{\partial^4 u}{\partial t^4}(\tau) \right\|^2_{L_2(\Omega)} d\tau \tag{8.3.3}$$

Setting $e_n = u_n - U^n$ and subtracting the linearized version of (8.2.11) from (8.3.1), we get

$$(\rho\delta_t^2 e_n, V)_o + a(e_n, V) = (\rho\varepsilon_n, V)_o \quad \forall \; V \in S_h(\Omega) \tag{8.3.4}$$

We again identify an element $W^n \in S_h(\Omega)$ through the weighted $H^1(\Omega)$ projection introduced by Wheeler [48]

$$a(u_n - W^n, V) = 0, \quad \forall \; V \in S_h(\Omega) \tag{8.3.5}$$

We decompose e_n by letting $e_n = E_n + \mathcal{E}_n$ where $E_n = u_n - W^n$ and $\mathcal{E}_n = W^n - U^n$. In addition we define certain auxiliary variables by

$$u_{n+1/2} = \frac{1}{2}(u_{n+1} + u_n)$$

$$\delta_t u_{n+1/2} = \frac{u_{n+1} - u_n}{\Delta t} \tag{8.3.6}$$

$$\delta_{t_{n+1/2}}(X) = \frac{X|_{t=(n+1)\Delta t} - X|_{t=n\Delta t}}{\Delta t}$$

The behavior of the error component $\mathcal{E}_n$ is given by the following Theorem:

<u>Theorem 8.1</u>. If $\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$ and $\frac{\Delta t^2}{h^2} \leq \frac{2\rho}{C'C^{*2}}$, then there exist positive constants C_1, C_2, not depending on the discretization parameters, such that

$$\|\delta_t \mathcal{E}\|_{\tilde{L}_\infty(L_2(\Omega))} + C_1 \|\mathcal{E}\|_{\tilde{L}_\infty(H^1(\Omega))}$$

$$\leq C_2 \{\|\mathcal{E}_0\|_{H^1(\Omega)} + \|\mathcal{E}_1\|_{H^1(\Omega)} + \|\delta_t \mathcal{E}_{\frac{1}{2}}\|_0 \tag{8.3.7}$$

$$+ \left\|\frac{\partial^2 E}{\partial t^2}\right\|_{L_2(L_2(\Omega))} + \Delta t^2 \left\|\frac{\partial^4 u}{\partial t^4}\right\|_{L_2(L_2(\Omega))}\}$$

where

$$\|\delta_t \mathcal{E}\|_{\tilde{L}_\infty(L_2(\Omega))} = \sup_{0 \leq n \leq N} \|\delta_t \mathcal{E}_{n-\frac{1}{2}}\|_0 \tag{8.3.8}$$

<u>Proof</u>: If follows from the decomposition of e_n and (8.3.4) that

$$(\rho\delta_t^2 E_n, V)_o + a(E_n, V) = -(\rho\delta_t^2 \mathcal{E}_n, V)_o$$

$$- a(\mathcal{E}_n, V) + (\rho\varepsilon_n, V)_o \quad \forall \quad V \in S_h(\Omega) \tag{8.3.9}$$

Now let $V = \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}}$; then

$$(\rho\delta_t^2 E_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})_o + a(E_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})$$

$$= -(\rho\delta_t^2 \mathrm{E}_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})_o \tag{8.3.10}$$

$$- a(\mathrm{E}_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}}) + (\rho\varepsilon_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})_o$$

But

$$a(E_n, \delta_t E_{n+\frac{1}{2}}) = -\frac{\Delta t}{2} a(\delta_t E_{n+\frac{1}{2}}, \delta_t E_{n+\frac{1}{2}})$$

$$+ \frac{1}{2} \delta_{t_{n+\frac{1}{2}}} a(E,E) \tag{8.3.11}$$

Similarly

$$a(E_n, \delta_t E_{n-\frac{1}{2}}) = \frac{\Delta t}{2} a(\delta_t E_{n+\frac{1}{2}}, \delta_t E_{n-\frac{1}{2}})$$

$$+ \frac{1}{2} \delta_{t_{n-\frac{1}{2}}} a(E,E) \tag{8.3.12}$$

and

$$(\rho\delta_t^2 E_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})_o = \frac{1}{\Delta t} [\|\delta_t E_{n+\frac{1}{2}}\|_o^2 \ \|\delta_t E_{n-\frac{1}{2}}\|_o^2] \tag{8.3.13}$$

Introducing these results into (8.3.11)

$$\frac{1}{\Delta t}\left[\| \delta_t E_{n+\frac{1}{2}} \|_o^2 - \| \delta_t E_{n-\frac{1}{2}} \|_o^2\right] - \frac{\Delta t}{2} a(\delta_t E_{n+\frac{1}{2}}, \delta_t E_{n+\frac{1}{2}})$$

$$+ \frac{\Delta t}{2} a(\delta_t E_{n-\frac{1}{2}}, \delta_t E_{n-\frac{1}{2}}) + \frac{1}{2} \delta_{t_{n+\frac{1}{2}}} a(E,E) + \frac{1}{2} \delta_{t_{n-\frac{1}{2}}} a(E,E)$$

$$= - (\rho\delta_t^2 E_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})_o - a(E_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})$$

$$+ (\rho\varepsilon_n, \delta_t E_{n+\frac{1}{2}} + \delta_t E_{n-\frac{1}{2}})_o \tag{8.3.14}$$

Eliminating the second term on the right hand side using (8.3.5), estimating the remaining terms on the right hand side using the Cauchy-Schwarz inequality and the elementary inequality E, multiplying by Δt, and summing from 1 to N-1,

$$\| \rho^{1/2}\delta_t E_{N-\frac{1}{2}} \|_o^2 - \| \rho^{1/2}\delta_t E_{\frac{1}{2}} \|_o^2 - \frac{\Delta t^2}{2} a(\delta_t E_{N-\frac{1}{2}}, \delta_t E_{N-\frac{1}{2}})$$

$$+ \frac{\Delta t^2}{2} a(\delta_t E_{\frac{1}{2}}, \delta_t E_{\frac{1}{2}}) + \frac{1}{2} a(E_n, E_n) + \frac{1}{2} a(E_{n-1}, E_{n-1})$$

$$- \frac{1}{2} a(E_1, E_1) - \frac{1}{2} a(E_o, E_o) \leq \Delta t \sum_{i=1}^{N-1} \{\kappa \| \rho^{1/2}\delta_t E_n \|_o^2$$

$$+ \xi \| \rho^{1/2}\varepsilon_n \|_o^2\} + \Delta t \sum_{i=1}^{N-1} \{\alpha \| \rho^{1/2}\delta_t E_{n+\frac{1}{2}} \|_o^2$$

$$+ \nu \| \rho^{1/2}\delta_t E_{n-\frac{1}{2}} \|_o^2\} \tag{8.3.15}$$

where κ, ξ, α, and ν are positive constants. Then using the Cauchy-Schwarz inequality, we conclude that there exists a positive constant C' such that

$$a(\delta_t E_{N-\frac{1}{2}}, \delta_t E_{N-\frac{1}{2}}) \leq C' \, \| \delta_t E_{N-\frac{1}{2}} \|^2_{H^1(\Omega)}$$

and using the inverse hypothesis on the subspace $S_h(\Omega)$ (8.2.8) we have

$$a(\delta_t E_{N-\frac{1}{2}}, \delta_t E_{N-\frac{1}{2}}) \leq \frac{C'C^{*2}}{h^2} \, \| \delta_t E_{N-\frac{1}{2}} \|^2_o \tag{8.3.16}$$

Introducing (8.3.14) into (8.3.13), using (8.3.4), and applying the Cauchy-Schwarz inequality.

$$\begin{aligned}
(1 - \frac{C'C^{*2}}{2\rho} \frac{\Delta t^2}{h^2}) \, \| \rho^{1/2} \delta_t E_{N-\frac{1}{2}} \|^2_o + \frac{\mu}{2} \, \| E_N \|^2_{H^1(\Omega)} \\
+ \frac{\mu}{2} \, \| E_{N-1} \|^2_{H^1(\Omega)} \leq \frac{C'}{2} \, \| E_o \|^2_{H^1(\Omega)} + \frac{C'}{2} \| E_1 \|^2_{H^1(\Omega)} \\
+ \| \rho^{1/2} \delta_t E_{\frac{1}{2}} \|^2_o \\
+ \Delta t \sum_{n=1}^{N-1} \{ \kappa \| \rho^{1/2} \delta_t E_n \|^2_o + \xi \| \rho^{1/2} \varepsilon_n \|^2_o \} \\
+ \Delta t \sum_{n=1}^{N-1} \{ \alpha \| \rho^{1/2} \delta^2_t E_{n+\frac{1}{2}} \|^2_o + \nu \| \rho^{1/2} \delta_t E_{n-\frac{1}{2}} \|^2_o \}
\end{aligned} \tag{8.3.17}$$

As a <u>condition</u> <u>of</u> <u>stability</u> we require that

$$(1 - \frac{C'C^{*2}}{2\rho} \frac{\Delta t^2}{h^2}) = C'' \geq 0 \tag{8.3.18}$$

which, as expected, places a constraint on the permissible values of the discretization parameters.

Finally, applying the discrete Gronwall inequality (Lemma 6.2), and the inequality

$$\Delta t \sum_{n=1}^{N-1} \| \delta_t^2 V_n \| \leq \| \frac{\partial^2 V}{\partial t^2} \|_{L_2(L_2(\Omega))} \tag{8.3.19}$$

we obtain (8.3.8). ∎

Then using Theorem 8.1, Lemma 6.1, and the triangle inequality, we obtain the final error estimate:

<u>Theorem 8.2</u>: If u, $u_t \in L_\infty(H^{k+1}(\Omega))$, $\frac{\partial^2 u}{\partial t^2} \in L_2(H^{k+1}(\Omega))$, $\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$, *and* $\frac{\Delta t^2}{h^2} \leq \frac{2\rho}{C'C^{*2}}$, there exist positive constants C_3 and C_4 such that

$$\| \delta_t e \|_{\tilde{L}_\infty(L^2(\Omega))} + C_3 \| e \|_{\tilde{L}_\infty(L_2(\Omega))}$$

$$\leq C_4 \{ \| e_o \|_{H^1(\Omega)} + \| e_1 \|_{H^1(\Omega)} + \| \delta_t e_{1/2} \|_o$$

$$+ h^{k+1} \| u \|_{L_\infty(H^{k+1}(\Omega))} + h^{k+1} \| \frac{\partial u}{\partial t} \|_{L_\infty(H^{k+1}(\Omega))}$$

$$+ h^{k+1} \| u \|_{L_2(H^{k+1}(\Omega))} + \Delta t^2 \| \frac{\partial^4 u}{\partial t^4} \|_{L_2(L_2(\Omega))} \}$$ ∎

<u>VIII.4 The Nonlinear Central Difference Approximation</u>. In this section, a priori bounds for the error involved in approximating (8.2.2) with

(8.2.11) are established. Initially we evaluate (8.2.2) at $t = n\Delta t$ and set $v = V$. The effect of the temporal approximation is determined by adding $(\rho\delta_t^2 u_n, V)_o$ to each side of the equation.

$$(\rho\delta_t^2 u_n, V)_o + a(u_n, u_n, V) = (f, V)_o + (\varepsilon_n, V)_o$$

$$\forall \quad V \in S_h(\Omega) \tag{8.4.1}$$

We use a decomposition of the error of the approximation described by (7.3.6). Let $W^n \in S_h(\Omega)$ be defined by the nonlinear energy projection introduced by Wheeler [48]:

$$a(u_n, u_n - W_n, V) = 0 \quad \forall \quad V \in S_h(\Omega) \tag{8.4.2}$$

Then subtracting (8.2.11) from (8.4.1), we have

$$(\rho\delta_t^2 e_n, V)_o + a(u_n, u_n, V) - a(U_n, U_n, V) = (\rho\varepsilon_n, V)_o$$

$$\forall \quad V \in S_h(\Omega) \tag{8.4.3}$$

Now we introduce certain conditions which for the nonlinear central difference scheme turn out to be sufficient for convergence.

I. <u>The Stability Condition</u>

$$\frac{\Delta t^2}{h^2} \leq \frac{2\rho}{M_2 C^{*2}} \tag{8.4.4}$$

II. The Response Condition

Let

$$L_n = M_3 \left\| \frac{U^{n+1} - U^{n-1}}{2\Delta t} \right\|_{L_\infty(\Omega)} \tag{8.4.5}$$

$$Q_n = \frac{\eta M_3^2 \xi^2 C^{*2}}{2} \sup_{1 \le i \le n} \left\| \frac{\partial W^n}{\partial x_i} \right\|^2_{L_\infty(\Omega)} \tag{8.4.6}$$

where η, ξ are positive constants. Then we require

$$\frac{\mu}{2}\left(\|E_N\|^2_{H^1(\Omega)} + \|E_{N-1}\|^2_{H^1(\Omega)} - \sum_{n=1}^{N-2} \Delta t L_n \|E_n\|^2_{H^1(\Omega)} \right.$$

$$- \sum_{n=1}^{N-1} \frac{\Delta t}{h^2} Q_n \|E_n\|^2_{H^1(\Omega)} \ge \Phi \|E_N\|^2_{H^1(\Omega)} \tag{8.4.7}$$

where $\Phi > 0$.

The behavior of the error component E_n is given in the following theorem:

Theorem 8.3. If $\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$ and the stability condition I and the response condition II are satisfied, then there exist positive constants C_1 and C_2, not depending on the discretization parameter such that

$$\|\delta_t E\|_{\tilde{L}_\infty(L_2(\Omega))} + C_1 \|E\|_{\tilde{L}_\infty(H^1(\Omega))}$$

$$\le C_1 \{ \|E_0\|_{H^1(\Omega)} + \|E_1\|_{H^1(\Omega)} + \|\delta_t E_{1/2}\|_0$$

$$+ \frac{1}{h} \|E\|_{L_\infty(L_2(\Omega))} + \|\frac{\partial^2 E}{\partial t^2}\|_{L_2(L_2(\Omega))}$$

$$+ \Delta t^2 \|\frac{\partial^4 u}{\partial t^4}\|_{L_2(L_2(\Omega))} \} \tag{8.4.8}$$

Proof: Decomposing e_n in (8.4.3).

$$(\rho\delta_t^2 E_n, V)_o + a(u_n, u_n, V) - a(U_n, U_n, V)$$

$$\leq - (\rho\delta_t^2 E_n, V) + (\rho\varepsilon_n, V)_o \quad \forall \quad V \in S_h(\Omega) \tag{8.4.9}$$

But

$$a(u_n, u_n, V) - a(U_n, U_n, V)$$

$$= a(u_n, E_n, V) + ((C^2(\underset{\sim}{x}, t, u_n) - C^2(\underset{\sim}{x}, t, W^n))\underset{\sim}{\nabla} W^n, \underset{\sim}{\nabla} V)_o$$

$$+ ((C^2(\underset{\sim}{x}, t, W^n) - C^2(\underset{\sim}{x}, t, U^n))\underset{\sim}{\nabla} W^n, \underset{\sim}{\nabla} V)_o + a(U_n, E_n, V) \tag{8.4.10}$$

Introducing (.4.10) into (.4.9) and setting $V = \delta_t E_{n+1/2} + \delta_t E_{n-1/2}$.

$$(\rho\delta_t^2 E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_o + a(U^n, E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})$$

$$+ ((C^2(\underset{\sim}{x}, t, W^n) - C^2(\underset{\sim}{x}, t, U^n))\underset{\sim}{\nabla} W^n, \underset{\sim}{\nabla}(\delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_o$$

$$= - (\rho\delta_t^2 E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_o - a(u_n, E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})$$

$$- ((C^2(\underset{\sim}{x}, t, u_n) - C^2(\underset{\sim}{x}, t, W^n))\underset{\sim}{\nabla} W^n, \underset{\sim}{\nabla}(\delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_o$$

$$+ (\rho\varepsilon_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_o \tag{8.4.11}$$

We next repeat steps used to obtain (8.3.13) to simplify the first term on the left and use (8.4.2) to eliminate the second term on the right.

$$\frac{1}{\Delta t}\,[\|\delta_t E_{n+1/2}\|_o - \|\delta_t E_{n-1/2}\|_o] + a(U^n, E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})$$

$$= -(\rho\delta_t^2 E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_o$$

$$-((C^2(\underset{\sim}{x},t,W^n) - C^2(\underset{\sim}{x},t,U^n))\underset{\sim}{\nabla}W^n, \underset{\sim}{\nabla}(\delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_o$$

$$-((C^2(\underset{\sim}{x},t,u_n) - C^2(\underset{\sim}{x},t,W^n))\underset{\sim}{\nabla}W^n, \underset{\sim}{\nabla}(\delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_o$$

$$+ (\rho\varepsilon_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_o$$

In order to estimate the terms in (8.4.12) we introduce certain useful relationships

$$a(U^n, E_n \delta_t E_{n+1/2}) = -\frac{\Delta t}{2}\, a(U^n, \delta_t E_{n+1/2}, \delta_t E_{n+1/2}) + \frac{1}{2}\,\delta_{t_{n+1/2}}\, a(U^n, E, E)$$

and

$$a(U^n, E_n, \delta_t E_{n-1/2}) = \frac{\Delta t}{2}\, a(U^n, \delta_t E_{n-1/2}, \delta_t E_{n-1/2}) + \frac{1}{2}\,\delta_{t_{n-1/2}}\, a(U^n, E, E)$$

Then

$$\Delta t \sum_{n=1}^{N-1} [a(U^n, E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}]$$

$$= -\frac{\Delta t^2}{2}\, a(U^{N-1}, \delta_t E_{N-1/2}, \delta_t E_{N-1/2}) + \frac{\Delta t^2}{2}\, a(U^o, \delta_t E_{1/2}, \delta_t E_{1/2})$$

$$+ \frac{\Delta t^2}{2} \sum_{n=1}^{N-1} [a(U^n,\delta_t E_{n-1/2},\delta_t E_{n-1/2}) - a(U^{n-1},\delta_t E_{n-1/2},\delta_t E_{n-1/2})]$$

$$+ \frac{1}{2} a(U^{N-1},E_N,E_N) + \frac{1}{2} a(U^{N-2},E_{N-1},E_{N-1}) - \frac{1}{2} a(U^2,E_1,E_1)$$

$$- \frac{1}{2} a(U^1,E_o,E_o) + \frac{1}{2} \sum_{n=1}^{N-2} [a(U^{n-1},E_n,E_n) - a(U^{n+1},E_n,E_n)] \quad (8.4.13)$$

Using $(8.2.4)_3$ and the Holder inequality, we obtain

$$\left| \frac{\Delta t^2}{2} \sum_{n=1}^{N-1} [a(U^n,\delta_t E_{n-1/2},\delta_t E_{n-1/2}) - a(U^{n-1},\delta_t E_{n-1/2},\delta_t E_{n-1/2}] \right|$$

$$\leq \sum_{n=1}^{N-1} \frac{\Delta t^3}{h^2} K_n \| \delta_t E_{n-1/2} \|_o^2$$

where

$$K_n = M_3 C^{*2} \| \frac{|U^n - U^{n-1}|}{\Delta t} \|_{L_\infty(\Omega)}$$

and

$$|\frac{1}{2} \sum_{n=1}^{N-2} [a(U^{n-1},E_n,E_n) - a(U^{n+1},E_n,E_n)]|$$

$$\leq \sum_{n=1}^{N-2} \Delta t L_n \| E_n \|^2_{H^1(\Omega)}$$

In addition, the inverse assumption (8.2.8)

$$\frac{\Delta t^2}{2}\, a(U^{N-1}, \delta_t E_{N-1/2}, \delta_t E_{N-1/2}) \le \frac{M_2 C^{*2}}{2\rho}\ \frac{\Delta t^2}{h^2}\ \| \rho^{1/2} \delta_t E_{N-1/2} \|_o^2$$

Thus from (8.4.13)

$$\Delta t \sum_{n=1}^{N-1} [a(U^n, E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})] \ge$$

$$- \frac{M_2 C^{*2}}{2\rho}\ \frac{(\Delta t)^2}{h^2}\ \| \rho^{1/2} \delta_t E_{N-1/2} \|_o^2 + \frac{\Delta t^2}{2}\, a(U^o, \delta_t E_{1/2}, \delta_t E_{1/2})$$

$$- \sum_{n=1}^{N-1} \frac{\Delta t^3}{h^2}\ K_n \| \delta_t E_{n-1/2} \| + \frac{1}{2}\, a(U^{N-1}, E_N, E_N)$$

$$+ \frac{1}{2}\, a(U^{N-2}, E_{N-1}, E_{N-1}) - \frac{1}{2}\, a(U^2, E_1, E_1) - \frac{1}{2}\, a(U^1, E_o, E_o)$$

$$- \sum_{n=1}^{N-1} \Delta t L_n \| E_n \|^2_{H^1(\Omega)} \qquad (8.4.14)$$

Similarly using the Holder inequality, the inverse assumption (8.4.8) and the imbedding result $\|E_n\|_{L_2(\Omega)} \le \xi \|E_n\|_{H^1(\Omega)}$ [3], we conclude that there exists a positive constant η such that

$$\Delta t \sum_{n=1}^{N-1} ((C^2(\underset{\sim}{x}, t, W^n) - C^2(\underset{\sim}{x}, t, U^n)) \underset{\sim}{\nabla} W^n, \underset{\sim}{\nabla}(\delta_t E_{n+1/2} + \delta_t E_{n-1/2}))_o$$

$$\le \sum_{n=1}^{N-1} \frac{\Delta t}{h^2}\ Q_n \| E_n \|^2_{H^1(\Omega)} + \frac{\Delta t}{2} \sum_{n=1}^{N-1} [\| \delta_t E_{n+1/2} \|_o^2$$

$$+ \| \delta_t E_{n-1/2} \|_o^2] \qquad (8.4.15)$$

where Q_n is defined by (8.4.6). In addition

$$\Delta t \sum_{n=1}^{N-1} ((C^2(\underset{\sim}{x},t,u_n) - C^2(\underset{\sim}{x},t,W^n))\underset{\sim}{\nabla} W^n, \underset{\sim}{\nabla}(\delta_t E_{n+1/2} + \delta_t E_{n+1/2}))_o$$

$$\leq \sum_{n=1}^{N-1} \frac{\Delta t}{h^2} \delta_n \| E_n \|_o^2 + \frac{\Delta t}{2\eta} \sum_{n=1}^{N-1} [\| \delta_t E_{n+1/2} \|_o^2 + \| \delta_t E_{n-1/2} \|_o^2] \tag{8.4.16}$$

where

$$\delta_n = \frac{\eta M_3^2 C^{*2}}{2} \sup_{1 \leq i \leq n} \| \frac{\partial}{\partial x_i} W^n \|_{L_\infty(\Omega)}^2 \tag{8.4.17}$$

Multiplying (8.4.12) by Δt, summing from 1 to N-1, and using

$$(1 - \frac{M_2 C^{*2}}{2\rho} \frac{\Delta t^2}{h^2}) \| \rho^{1/2} \delta_t E_{N-1/2} \|_o^2$$

$$+ \frac{1}{2} a(U^{N-1}, E_N, E_N) + \frac{1}{2} a(U^{n-2}, E_{N-1}, E_{N-1}) - \sum_{n=1}^{N-2} \Delta t L_n \| E_n \|_{H^1(\Omega)}^2$$

$$- \sum_{n=1}^{N-1} \frac{\Delta t}{h^2} Q_n \| E_n \|_{H^1(\Omega)}^2 \leq \| \rho^{1/2} \delta_t E_{1/2} \|_o^2 + \frac{1}{2} a(U^2, E_1, E_1)$$

$$+ \frac{1}{2} a(U^1, E_o, E_o) + \Delta t \sum_{n=1}^{N-1} \frac{S_n}{h^2} \| E_n \|_o^2$$

$$- \Delta t \sum_{n=1}^{N-1} (\rho \delta_t^2 E_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2}) + \frac{\Delta t}{\eta} \sum_{n=1}^{N-1} [\| \delta_t E_{n+1/2} \|_o^2$$

$$+ \;\|\delta_t E_{n-1/2}\|_o^2] + \Delta t \sum_{n=1}^{N-1} \frac{\Delta t^2}{h^2} K_n \|\delta_t E_{n-1/2}\|_o^2$$

$$+ (\rho\varepsilon_n, \delta_t E_{n+1/2} + \delta_t E_{n-1/2})_o \tag{8.4.18}$$

Now using the stability condition (I) and the response condition (II), estimating the terms on the right hand side using the Cauchy-Schwarz inequality and inequality E, applying the Gronwall inequality (Lemma 6.3) and using (8.4.4), we obtain the result (8.4.8).

Now using Theorem .3, Lemma .2, and the triangle inequality, we obtain the final error estimate. ■

Theorem 8.4. If $u, \frac{\partial u}{\partial t} \in L_\infty(H^{k+1}(\Omega))$, $\frac{\partial^2 u}{\partial t^2} \in L_2(H^{k+1}(\Omega))$, $\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$, and the stability condition I and the response condition II are satisfied, then there exist positive constants C_3 and C_4 such that

$$\|\delta_t e\|_{\tilde{L}_\infty(L_2(\Omega))} + C_3\|e\|_{\tilde{L}_\infty(L_2(\Omega))}$$

$$\leq C_4\{\|e_o\|_{H^1(\Omega)} + \|e_1\|_{H^1(\Omega)}$$

$$+ \|\delta_t e_{1/2}\|_o + h^k\|u\|_{L_\infty(H^{k+1}(\Omega))} + h^{k+1}\left\|\frac{\partial u}{\partial t}\right\|_{L_\infty(H^{k+1}(\Omega))}$$

$$+ h^{k+1}\left\|\frac{\partial^2 u}{\partial t^2}\right\|_{L_2(H^{k+1}(\Omega))} + \Delta t^2\left\|\frac{\partial^4 u}{\partial t^4}\right\|_{L_2(L_2(\Omega))}\} \tag{8.4.19}$$

We can make certain comments concerning these results.

AD-A031 238 TEXAS UNIV AT AUSTIN DEPT OF AEROSPACE ENGINEERING AN--ETC F/G 20/11
A STUDY OF CONVERGENCE AND STABILITY OF FINITE ELEMENT APPROXIM--ETC(U)
AUG 76 J T ODEN, L C WELLFORD, C T REDDY DAAG29-76-G-0022

UNCLASSIFIED ARO-11860.9-M NL

4 OF 4
ADA031238

END
DATE
FILMED
11 - 76

(i) A version of the Courant-Friedricks-Lewy stability criteria [4] is required for the numerical stability of the linear central difference scheme. The constants in the C.-F.-L. stability criteria are determined from the inverse hypotheses for the subspace $S_h(\Omega)$ and the continuity property of the bilinear form a(u,v).

(ii) To obtain numerical stability for the nonlinear central difference approximation, we require that the C.F.L. stability criteria be satisfied and that the amplitude of the response be small. The constant in the C.F.L. stability criteria is determined from the inverse hypothesis on the subspace of $S_h(\Omega)$ and the bound on the wave speed squared defined in (8.2.4). The limiting value of the response is determined from the response condition.

(iii) The rate of convergence in the nonlinear central difference is lower by one power in the spatial discretization parameter h as compared to the linear central difference approximation. This means that the nonlinear central difference approximation does not obtain the optimum spatial rate of convergence of the subspace $S_h(\Omega)$.

(iv) The error in the initial data is critical in the central difference approximation. In both the linear and nonlinear estimates the error in the initial data at the first and second time points appears. Of course, the error at the second time point occurs because the scheme is not self-starting. In particular, the error in the displacement in the $H^1(\Omega)$ norm occurs at the first two time points, and the error in the discrete velocity in the $L_2(\Omega)$ norm occurs at the midpoint between time point one and two. We note that the error in the initial data occurs in the $H^1(\Omega)$ norm even though the error estimate is in a

weaker norm (the $L_2(\Omega)$ norm). Thus we must specify the initial data very accurately and in addition use a very accurate starting procedure.

(v) In order to obtain convergence of the linear and nonlinear central difference approximation, we require the following regularity of the exact solution:

$$u \in L_\infty(H^{k+1}(\Omega))$$

$$\frac{\partial u}{\partial t} \in L_\infty(H^{k+1}(\Omega))$$

$$\frac{\partial^2 u}{\partial t^2} \in L_2(H^{k+1}(\Omega))$$

$$\frac{\partial^4 u}{\partial t^4} \in L_2(L_2(\Omega))$$

CHAPTER IX

SUMMARY AND CONCLUSIONS

IX.1 Contributions and Results. A general theory of nonconforming generalized variational approximations is developed. These approximations which we have identified as generalized Galerkin methods use discontinuous trial functions, and thus are valid for problems with irregular data and corresponding irregular solutions. This method is shown to be an approximation to the solution of the original problem in its free boundary form. The generalized Galerkin method is formulated for the problem of wave and shock propagation in certain nonlinear elastic solids.

The accuracy and convergence of the generalized Galerkin method for shocks is analyzed. In this analysis we show as an intermediate result that the standard Galerkin method cannot converge to shock wave solution. We show that the new technique converges to shock wave solution but that at the shock it loses accuracy compared to the standard Galerkin method on shockless domains. A criteria for the numerical stability of the fully discretized generalized Galerkin method is developed. Stability is shown to be a much more complicated problem as compared to the normal Galerkin scheme (which is governed by the Courant-Fredricks-Lewy criteria.)

The generalized Galerkin shock procedure is shown to lead to shock fitting methods when implemented with finite-element interpolants. Numerical procedures are introduced for the calculation of the propagation of shock and acceleration waves using these procedures. It should be noted that the chief feature of this shock fitting scheme is that the position of the free boundary (shock) is itself one of the dependent variables of the problem. Techniques for calculating the reflection of waves are introduced, and numerical examples are presented.

A theory of the approximation of certain nonlinear elasticity problems by the finite-element method is developed. In particular we have considered here hyperelastic materials. The first Piola-Kirchhoff stress tensor is characterized physically and mathematically for these materials. The solvability and regularity of the original problem is discussed. The convergence and rate of convergence for the finite element approximation of these problems is determined.

A theory of parabolic regularization methods for shock smearing applications is developed. The accuracy and convergence of the parabolic regularization method are determined. Numerical stability criteria for the method are developed and are shown to be highly restrictive. The effect of the size of the regularizing parameter on the accuracy, convergence, and stability of the scheme is demonstrated.

The parabolic regularization method is implemented using finite element trial spaces. Numerical results in the problem of shock propagation are presented.

A theory of central difference approximations for shockless nonlinear dynamic response is presented. The accuracy and convergence of the method are determined. Numerical stability criteria for the

method are derived. It is shown that to insure numerical stability a condition limiting the level of the response must be imposed.

IX.2 Conclusions. The generalized Galerkin theory developed here is a new technique which is much more general than at first supposed. It appears to be a basic nonconforming method which is particularly useful for irregular problems (problems with point loads, interface problems, problems with irregular domains and corners, and, of course, shocks). The technique is particularly novel in that it uses discontinuous trial functions.

The generalized Galerkin technique in its formulation for problems of shock propagation leads to a shock fitting scheme. The position of the shcok wave appears as one of the dependent variables. The approximation of shock waves using this technique seems to be very effective. It is believed that this technique can vastly improve the potential of shock fitting schemes using noncharacteristic methods for one and two-dimensional problems.

The analysis of the nonlinear static elasticity problem represents the first substantial inroad into the theory of approximations of nonlinear elasticity problems. For the first time we see that theoretically as well as experimentally the behavior of nonlinear elasticity problems is significantly different as compared to linear problems. We conclude from this work that for very regular problems the nonlinear finite-element technique acts like a linear approximation while for irregular problems the nonlinear technique is significantly different. We conclude that if higher order terms are used in the constitutive law for stress for irregular problems, then the degree of the polynomial in the finite element approximation should be increased.

The parabolic regularization scheme introduced in this work appears to be effective in computing shock waves. The theoretical methods used to characterize the scheme mathematically seem to be very general and will serve as a model for the analysis of other shock smearing schemes. In particular the incorporation of a regularizing parameter in the approximation scheme seems to be effective theoretically and computationally. Theoretically this allows us to use the regularity theory of partial differential equations to bring the regularizing parameter into the error estimates and into the stability criteria. Computationally this allows us to adjust the number of elements over which the shock is smeared.

The theory of central difference methods for nonlinear shockless dynamic response allows us to determine the effect of the nonlinearity on the accuracy and stability.

IX.3 Future Research. It is believed that the following areas represent significant future directions for this research:

(i) Irregular elliptic problems. The generalized Galerkin method with discontinuous trial functions should be applied initially to one-dimensional elliptic problems with singularities, point leads, and interfaces. The true position of this method relative to other standard elliptic techniques such as the mixed and hybrid models would then be determined. Extensions should be made to two-dimensional problems.

(ii) *Free boundary problems*. The generalized Galerkin method should be applied to other problems which are irregular due to a moving free boundary. In particular, the Stefan problem and the deformation of

elastic, perfectly plastic solid are examples. Here the advantage of the generalized Galerkin method is clear. The position of the free boundary is one of the dependent variables of the problem and is computed explicitly.

(iii) Nonlinear elasticity in two dimensions. Extension of the theoretical results on accuracy and convergence of nonlinear elasticity problems to several dimensions should be carried out.

(iv) Additional regularizing schemes for shock waves. Determination of other regularizing schemes for shock waves. Comparison of these new schemes theoretically and computationally.

(v) Two dimensional shock propagation. The use of the generalized Galerkin shock fitting scheme in two dimensions should be investigated.

(vi) Hybrid method for shock propagation. The feasibility and potential of a more sophisticated model for the wave front should be investigated. In particular a hybrid model should be considered.

(vii) H^{-1} Galerkin procedures. The relationship between the generalized Galerkin models and the H^{-1} Galerkin procedures of Rachford and Wheeler [78] should be investigated. A model which is a cross between the H^{-1} method and the generalized Galerkin technique should be formulated for shocks.

(viii) Structural dynamic temporal operators. The standard structural dynamic temporal operators (Houbolt, New mark, Wilson, etc.) should be analyzed using the L_2 methods of Chapter VIII for linear and nonlinear problems. A comparative study should be initiated to determine the relative advantages of each.

B I B L I O G R A P H Y

1. Von Neumann, J., "Proposal and Analysis of a Numerical Method for the Treatment of Hydrodynamical Schocks," Nat. Defense and Research Comm., Report AM-551, 1944.

2. Lax, P., Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves, Society for Industrial and Applied Mathematics, Philadelphia, 1973.

3. Sobolev, S. L., Applications of Functional Analysis in Mathematical Physics, Vol. 7., Translations of Mathematical Monographs, American Mathematical Society, Providence, 1963.

4. Courant, R., Friedrichs, K., and Lewy, H., "The Partial Difference Equations of Mathematical Physics," IBM Journal, 1967, pp. 215-234 (Translation of the original article in Math. Ann., 100, 1928, pp. 32-51).

5. Von Neumann, J. and Richtmyer, R., "A Method for the Numerical Calculation of Hydrodynamical Shocks," Journal of Applied Physics, 21, 1950, pp. 232-257.

6. Landshoff, R., "A Numerical Method for Treating Fluid Flow in the Presence of Shocks," Los Almos Scientific Laboratory, Report Number LA-1930, 1955.

7. Rusanov, V. V., "Calculation of Interaction of Non-Steady Shock Waves with Obstacles," U.S.S.R. Journal of Computational Mathematics and Mathematical Physics, 1, 1961, pp. 267-279.

8. Lax, P. D., "Weak Solutions of Nonlinear Hyperbolic Equations and Their Numerical Computation," Communications on Pure and Applied Mathematics, 7, 1954, pp. 159-193.

9. Lax, P. D., and Wendroff, B., "Systems of Conservation Laws," Communications on Pure and Applied Mathematics, 13, 1960, pp. 217-237.

10. Lax, P. D., and Wendroff, B., "Difference Schemes with High Order of Accuracy for Solving Hyperbolic Equations," Communications on Pure and Applied Mathematics, 17, 1964, pp. 381-401.

11. Richtmyer, R., and Morton, K., Difference Methods for Initial Value Problems, Interscience, New York, 1967.

12. MacCormack, R., Numerical Solution of the Interaction of a Shock Wave with a Laminar Boundary Layer, Lecture Notes in Physics, No. 8, Springer-Verlag, Berlin, 1971.

13. Rusanov, V. V., "On Difference Schemes of Third Order Accuracy for Nonlinear Hyperbolic Systems," Journal of Computational Physics, 5, 1970, pp. 507-516.

14. Burstein, S. Z., and Mirin, A. A., "Third Order Accurate Methods for Hyperbolic Equations," Journal of Computational Physics, 5, 1970, pp. 547-571.

15. Godunov, S., "An Interesting Class of Quasilinear Systems," Dokl. Akad. Nauk. SSSR, 139, 1961, pp. 521-523.

16. Glimm, J., "Solutions in the Large for Nonlinear Hyperbolic Systems of Equations," Communications on Pure and Applied Mathematics, 18 1965, pp. 697-715.

17. Courant, R., Isaacson, E., and Rees, M., "On the Solution of Nonlinear Hyperbolic Differential Equations by Finite Differences," Communications on Pure and Applied Mathematics, 5, 1952, pp. 243-255.

18. Walsh, R. T., "Finite Difference Methods," Dynamic Response of Materials to Intense Implusive Loading, Ed. P. C. Chou and A. K. Hopkins, Air Force Materials Laboratory, 1973, pp. 363-403.

19. Lawerence, R. J. "WONDY III A-A Computer Program for One-Dimensional Wave Propagation," Sandia Laboratories, Report SC-DR-70-315, 1970.

20. Lawerence, R. J., and Mason, D. S., "WONDY IV - A Computer Program for One-Dimensional Wave Propagation with Rezoning," Sandia Laboratories, Report SC-RR-71-0284, 1971.

21. Moretti, G., "Floating Shock Fitting Techniques for Imbedded Shock in Unsteady Multidimensional Flows," Proceedings of the 1974 Heat Transfer and Fluid Mechanics Institute, Stanford University Press, 1974.

22. Moretti, G. "A Critical Analysis of Numerical Techniques: The Piston-Driven Inviscid Flow," Polytechnic Institute of Brooklyn (PIBAL) Report 69-25, 1969.

23. Moretti, G., "Experiments in Multi-Dimensional Floating Shock Fitting," PIBAL Report 73-18, 1973.

24. Moretti, G., "Thoughts and Afterthoughts About Shock Computations," PIBAL Report 72-37, 1972.

25. Moretti, G., "The Importance of Boundary Conditions in the Numerical Treatment of Hyperbolic Equations," PIBAL Report 68-34, 1968.

26. Moretti, G., "Inviscid Blunt Body Shock Layers-Two Dimensional Symmetric and Axisymmetric Flows," PIBAL Report 68-15, 1968.

27. Moretti, G., and Abbett, M., "A Time-Dependent Computational Method for Blunt Body Flows," A.I.A.A. Journal, 4, 1966, pp. 2136-2141.

28. Moretti, G., and Bleich, G., "Three Dimensional Flow About Blunt Bodies," A.I.A.A. Reprint 67-222, 1967.

29. Yu, N. J., and Seebass, R., "Computational Procedures for Mixed Equations with Shock Waves," Computational Methods in Nonlinear Mechanics, Texas Institute for Computational Mechanics, Austin, Texas, 1974, pp. 499-508.

30. Skoglund, V. J., and Cole, J. K., "Numerical Analysis of an Oblique Shock Wave and a Laminar Boundary Layer," Bureau of Engineering Research Report No. ME-22, University of New Mexico, 1966.

31. Taylor, T. P., "Computing Transient Gas Flows with Shock Waves," Physics of Fluids, 7, 1964, pp. 1713-1715.

32. Bellman, R., Cherry, I., and Wing, G. M., "A Note on the Numerical Integration of a Class of Nonlinear Hyperbolic Equations," Quarterly Journal of Applied Mathematics, 16, 1958, pp. 181-183.

33. Gary, J., "Numerical Computation of Hydrodynamics Flows which Contain a Shock," Report N.Y.O. 9603, Courant Inst. Math. Sci., New York University, 1962.

34. Xerikos, J., "A Time-Dependent Approach to the Numerical Solution of the Flow Field about an Axisymmetric Vechicle at Angle of Attack," Douglas Aircraft Co., Santa Monica, Calif., 1968.

35. Xerikos, J., and Anderson, W. A., "A Time-Dependent Approach to the Numerical Solution of the Flow Field About an Axisymmetric Vehicle at an Angle of Attack," NASA Cr-61982, 1968.

36. Weinbaum, S., "Rapid Expansion of a Supersonic Boundary Layer and Its Application to the Near Wake," A.I.A.A. Journal, 4, 1966, pp. 217-227.

37. Karpp, R., and Chou, P. C., "The Method of Characteristics," Dynamic Response of Materials to Intense Implusive Loading, Ed. P. C. Chou and A. K. Hopkins, Air Force Materials Laboratory, 1973.

38. Thomas, L. H., "Computation of One-Dimensional Compressible Flows Including Shocks," Communications on Pure and Applied Mathematics, 7, 1954, pp. 195-206.

39. Herrman, W., Hicks, D. L., and Young, E. G., "Attenuation of Elastic Plastic Stress Waves," Shock Waves and the Mechanical Properties of Solids, Syracuse University Press, 1971, pp. 23-63.

40. Hicks, D. L., and Holdridge, P. B., "The Conchas Wavecode," Sandia Laboratories, SC-RR-72-0451, 1972.

41. Oden, J. T., and Fost, R. B., "Convergence, Accuracy and Stability of Finite-Element Approximations of a Class of Nonlinear Hyperbolic Equations," International Journal of Numerical Methods in Engineering, 6, 1973, pp. 357-365.

42. Fost, R. B., Finite Element Analysis of Shock and Finite-Amplitude Waves in Hyperelastic Bodies at Finite Strain, Ph.D. Dissertation, Georgia Institute of Technology, 1974.

43. Fost, R. B., Oden, J. T., and Wellford, L. C. Jr., "A Finite Element Analysis of Shock and Finite-Amplitude Waves in One-Dimensional Hyperelastic Bodies at Finite Strain," International Journal of Solids and Structures, (to appear).

44. Oden, J. T., Key, J. E., and Fost, R. B., "A Note on the Analysis of Nonlinear Dynamics of Elastic Membranes by the Finite Element Method," Computers and Structures, 10, 1974, pp. 445-452.

45. Oden, J. T., "Approximations and Numerical Analysis of Finite Deformations of Elastic Solids," Nonlinear Elasticity, Ed. W. Dickey, Academic Press, New York, 1974, pp. 175-228.

46. Argyris, J. H., Dunne, P. C., Angelopoulos, T., and Bichat, B., "Finite Element Difficulties in Nonlinear and Incompressible Materials," Computational Methods in Nonlinear Mechanics, Texas Institute for Computational Mechanics, Austin, Texas, 1974, pp. 747-756.

47. Douglas, J., and Dupont, T., "Galerkin Methods for Parabolic Equations," S.I.A.M. Journal on Numerical Analysis, 7, 1970, pp. 575-626.

48. Wheeler, M. F., "A Priori L_2 Error Estimates for Galerkin Approximations to Parabolic Partial Differential Equations," S.I.A.M. Journal on Numerical Analysis, 10, 1973, pp. 723-759.

49. Dupont, T., "L^2-Estimates for Galerkin Methods for Second Order Hyperbolic Equations," S.I.A.M. Journal on Numerical Analysis, 10, 1973, pp. 880-889.

50. Douglas, J., and Dupont, T., "Alternating Direction Galerkin Methods on Rectangles," Proc. Symposium on Numerical Solution of Partial Differential Equations, II, Ed. B. Hubbard, Academic Press, 1971, pp. 133-214.

51. Dendy, J., "An Analysis of Some Galerkin Schemes for the Solution of Nonlinear Time-Dependent Problems," Los Almos Scientific Laboratories, LA-UR-73-1836, 1973.

52. Oden, J. T., "Approximations and numerical analysis of finite deformations of elastic solids," Nonlinear Elasticity, edited by R. W. Dickey, Academic Press, N. Y. (1973), pp. 175-228.

53. Treloar, L. R. G., "Stress-strain data for vulcanized rubber under various types of deformation," Trans. Faraday Soc., 40 (1944), p. 59.

54. Mooney, M., "A Theory of large elastic deformations," J. Appl'd. Phys., 11 (1940), pp. 582-592.

55. Biderman, V. L., "Calculation of rubber parts (in Russian)," Rascheti na Prochnost, Moscow (1958).

56. Isihara, A., Hashitsume, N., and Tatibana, M., "Statistical theory of rubber-like elasticity - IV (Two-dimensional stretching)," J. Chem. Phys., 19 (1951), pp. 1508-1512.

57. Kachurovskii, R., "Non-Linear Monotone Operators in Banach Spaces," Russian Math. Surveys, 23 (1968), pp. 117-165.

58. Glowinski, R., and Marrocco, A., "Sur l'approximation, par elements finis d'ordre 1, et la resolution, par penalisation d'une classe de problemes de Dirichlet non lineaires," Comptes Rendus Acad. Sci. Paris, Tome 278, Serie A, 1974, pp. 1649-1652.

59. Glowinski, R., and Marrocco, A., Private Communication.

60. Aubin, J. P., Approximation of Elliptic Boundary-Value Problems, Wiley Interscience, New York, 1972.

61. Oden, J. T., Finite Elements of Nonlinear Continua, McGraw-Hill, New York, 1972.

62. Ciarlet, P. G., and Raviart, P. A., "General Lagrange and Hermite Interpolation in $\mathbb{R}^n$ with Applications to Finite Element Methods," Archives for Rational Mechanics and Analysis, 46, 1972, pp. 177-199.

63. Strang, G., "Approximation in the Finite Element Method," Numerische Mathematik, 19, 1972, pp. 81-98.

64. Oden, J. T., and Reddy, C. T., "Finite Element Approximation of a Two-Point Boundary Value Problem in Nonlinear Elasticity", Journal of Elasticity (to appear).

65. Nitsche, J., "Ein kriterium für die quasi-optimalität des Ritzschen verfahrens," Numer. Math., 11 (1968), pp. 346-348.

66. Aubin, J. P., "Evaluation des erreurs de troncature des approximations des espaces de Sobolev," J. Math. Anal. Appl., 21 (1968), pp. 356-368.

67. Strang, G. and Fix, F., An analysis of the finite element method, Prentice-Hall, Englewood Cliffs, N.J., 1973.

68. Nečas, J., Les Methodes Directes en Theorie des Equations Elliptiques, Masson, Paris, 1967.

69. Tartar, L., "Interpolation non Lineaire et Regularite," Journal of Functional Analysis, 9 (1972), pp. 469-489.

70. Vainberg, M. M., Variational methods for the study of nonlinear operators (translated from the 1956 Russian edition by A. Feinstein), Holden-Day, San Francisco (1964).

71. Raviart, P. A., "Sur l'Approximation de Certaines Equations d'Evolution Lineaires et Non Lineaires," Journal de Math. Pures et Appl., 46, 1967, pp. 11-183.

72. Bellman, R., Stability Theory of Differential Equations, McGraw-Hill, 1952.

73. Lees, M., "A Priori Estimates for the Solution of Difference Approximations to Parabolic Partial Differential Equations," Duke Mathematical Journal, 27, 1960, pp. 297-311.

74. Truesdell, C., and Toupin, R., "The Classical Field Theories," Encyclopedia of Physics, Vol. III/1, Ed. S. Flugge, Springer-Verlag, 1960, pp. 226-793.

75. Chen, P. J., "Growth and Decay of Waves in Solids," Encyclopedia of Physics, VI a/3 - Mechanics of Solids, III, Ed. S. Flugge, Springer-Verlag, 1973, pp. 303-402.

76. Babuška, I., and Aziz, A. K., "Survey lectures on the mathematical foundations of the finite element method," The mathematical foundations of the finite element method with applications to partial differential equations, edited by A. K. Aziz, Academic Press, N.Y. (1972) pp. 3-345.

77. Fujii, H., "Finite Element Schemes: Stability and Convergence," Advances in Computational Methods in Structural Mechanics and Design, University of Alabama at Huntsville Press, 1972, pp. 201-218.

78. Rachford, H., and Wheeler, M., "An H^{-1} Galerkin Procedure for the Two-Point Boundary-Value Problem," In review.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER P/11860-M	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) A Study of Convergence and Stability of Finite Element Approximation of Shock and Acceleration Waves in Nonlinear Materials.		5. TYPE OF REPORT & PERIOD COVERED Final Report 9/1/75-8/31/76
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) J. /Oden, L. C. /Wellford, Jr. [redacted] C. T. /Reddy Tinsley		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-76-G-0022 NEW
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of the Army U. S. Army Research Office P. O. Box 12211, Research Triangle Park, N.C. 27709		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 1T161102BH57-05 Mathematics
11. CONTROLLING OFFICE NAME AND ADDRESS Dr. Jagdish Chandra Mathematics Division ARO, P. O. Box 12211, Research Triangle Park, N.C.		12. REPORT DATE (11) Aug[redacted] 1976
		13. NUMBER OF PAGES 298
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (9) Final rept. 1 Sep 75-31 Aug 76		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution Unlimited (18) ARO (19) 11860.9-M (12) 305p.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(16) DA-1-T-161102-BH-57

18. SUPPLEMENTARY NOTES

(17) 1-T-161102-BH-5705

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Finite Element Methods, Shock Waves, Nonlinear Hyperbolic Equations, Nonlinear Elasticity, Shock Fitting Methods

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes the recent work on the development of discontinuous finite element methods for the analysis of shock waves in nonlinear elastic materials. A class of one-dimensional finite elements is introduced in which the local interpolation functions consist of the usual piecewise linear functions and some additional functions which have discontinuities. In this way it is possible to model the local displacement field in terms of the values of the displacement at each node and two additional terms in which the shock strength and the location of the shock within an element are used as parameters.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

402 465 (over)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LB

20. (Continuation) The corresponding variational formulation contains the required jump conditions. For a specific class of material a priori error estimates are derived and the scheme is implemented and applied to a number of representative examples.